

17 JUN 2005

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2005 年 3 月 17 日 (17.03.2005)

PCT

(10) 国際公開番号  
WO 2005/024788 A1

- (51) 国際特許分類<sup>7</sup>: G10L 21/02  
 (21) 国際出願番号: PCT/JP2004/012629  
 (22) 国際出願日: 2004 年 9 月 1 日 (01.09.2004)  
 (25) 国際出願の言語: 日本語  
 (26) 国際公開の言語: 日本語  
 (30) 優先権データ:  
     特願2003-309720 2003 年 9 月 2 日 (02.09.2003) JP  
     特願2004-195818 2004 年 7 月 1 日 (01.07.2004) JP  
     特願2004-195867 2004 年 7 月 1 日 (01.07.2004) JP  
 (71) 出願人 (米国を除く全ての指定国について): 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町二丁目 3 番 1 号 Tokyo (JP).  
 (72) 発明者; および  
 (75) 発明者/出願人 (米国についてのみ): 荒木 章子

(ARAKI, Shoko) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 NTT 知的財産センタ内 Tokyo (JP). 澤田 宏 (SAWADA, Hiroshi) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 NTT 知的財産センタ内 Tokyo (JP). 牧野 昭二 (MAKINO, Shoji) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 NTT 知的財産センタ内 Tokyo (JP). 向井 良 (MUKAI, Ryo) [JP/JP]; 〒1808585 東京都武蔵野市緑町三丁目 9 番 1 1 号 NTT 知的財産センタ内 Tokyo (JP).

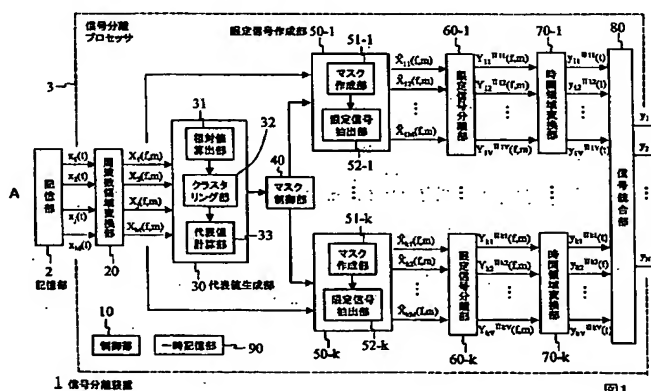
(74) 代理人: 草野 卓, 外 (KUSANO, Takashi et al.); 〒1600022 東京都新宿区新宿三丁目 1 番 2 2 号 新宿 NSOビル4階 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU,

[続葉有]

(54) Title: SIGNAL SEPARATION METHOD, SIGNAL SEPARATION DEVICE, SIGNAL SEPARATION PROGRAM, AND RECORDING MEDIUM

(54) 発明の名称: 信号分離方法、信号分離装置、信号分離プログラム及び記録媒体



1... SIGNAL SEPARATION DEVICE  
 2... STORAGE UNIT  
 A... STORAGE UNIT  
 3... SIGNAL SEPARATION PROCESSOR  
 20... FREQUENCY REGION CONVERSION UNIT  
 10... CONTROL UNIT  
 90... TEMPORARY STORAGE UNIT  
 31... RELATIVE VALUE CALCULATION UNIT  
 32... CLUSTERING UNIT  
 33... RELATIVE VALUE CALCULATION UNIT  
 30... REPRESENTATIVE VALUE GENERATION UNIT  
 40... MASK CONTROL UNIT

50-1... LIMIT SIGNAL CREATION UNIT  
 51-1... MASK CREATION UNIT  
 52-1... LIMIT SIGNAL EXTRACTION UNIT  
 51-k... MASK CREATION UNIT  
 52-k... LIMIT SIGNAL EXTRACTION UNIT  
 60-1... LIMIT SIGNAL SEPARATION UNIT  
 60-k... LIMIT SIGNAL SEPARATION UNIT  
 70-1... TIME REGION CONVERSION UNIT  
 70-k... TIME REGION CONVERSION UNIT  
 80... SIGNAL INTEGRATION UNIT

(57) Abstract: When the number of signal sources  $N$  and the number of sensors  $M$  is in the relationship of  $N > M$ , a mixed signal can be separated with a high quality. Firstly, a value of an observed signal which has been observed in the sensor is converted in a signal value of a frequency region. By using the signal value in the frequency region, a relative

[続葉有]



ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告書

- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE,

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

value of the observed value between the sensors is calculated for each frequency. Next, each relative value is clustered into N classes and a representative value of each class is calculated. By using the representative values, a mask is created for extracting a value of a signal transmitted from the V ( $1 \leq V \leq M$ ) signal sources. By using this mask, a value of the signal transmitted from the V signal sources is extracted from the signal value of the aforementioned frequency region. After this, if  $V = 1$ , the limit signal is directly outputted as a separation signal and if  $V \geq 2$ , ICA or the like is applied to the limit signal so as to obtain a separation signal.

(57) 要約: 信号源の数Nとセンサの数Mが $N > M$ の関係にある場合に、混合信号を高い品質で分離する。まず、センサにおいて観測された観測信号の値を周波数領域の信号値に変換し、周波数領域の信号値を用い、センサ間における観測値の相対値を、各周波数において算出する。次に各相対値をN個のクラスにクラスタリングし、各クラスの代表値を算出する。そして、これらの代表値を用い、周波数領域の信号値から、V ( $1 \leq V \leq M$ ) 個の信号源から発せられた信号の値を抽出するためのマスクを作成し、このマスクを用い、前記周波数領域の信号値から、V個の信号源から発せられた信号の値を抽出する。その後、 $V = 1$ であれば限定信号をそのまま分離信号として出力し、 $V \geq 2$ であればこの限定信号にICA等を適用し、分離信号を得る。

## 明 細 書

信号分離方法、信号分離装置、信号分離プログラム及び記録媒体  
技術分野

- [0001] 本発明は、信号処理の技術分野に関し、特に、必要である源信号(目的信号)のみを直接観測することができず、目的信号に他の信号が重畳されて観測されるという状況において目的信号を推定する信号分離方法、信号分離装置、信号分離プログラム及びそれを格納した記録媒体に関する。

## 背景技術

- [0002] 従来から、複数の源信号(音声信号等)が混合された混合信号を用い、源信号や混合過程の知識を用いることなく、混合前の源信号を分離・抽出するブラインド信号分離(BSS: Blind Source Separation)技術が知られている。

図27Aは、このブラインド信号分離技術を概念的に例示したブロック図である。

この図に例示するように、ブラインド信号分離では、複数(この例ではN個)の信号源701から発せられた源信号 $s_i$  ( $i=1, \dots, N$ )が混合し、複数(この例ではM個)のセンサ702で観測される状況下において、その観測信号 $x_j$  ( $j=1, \dots, M$ )のみから、源信号と推測される分離信号 $y_k$  ( $k=1, \dots, N$ )を取り出す。ここで、信号源701から発せられた源信号 $s_i$ が混合し、センサ702で観測されるまでの過程を「混合過程」と呼び、センサ702の観測結果から分離信号を取り出す過程を「分離過程」と呼ぶ。

- [0003] はじめに、観測される信号及び分離問題を定式化する。

〔実環境での混合信号(観測信号)のモデル〕

まず、混合過程についてモデル化する。

Nを信号源701の個数、Mをセンサ702の個数、 $s_i$ をi番目の信号源701(信号源i)から発せられた信号(源信号)、 $h_{ji}$ を信号源iからj番目のセンサ702(センサj)までのインパルス応答とする。この場合、センサjで観測される信号 $x_j$ は、これら源信号 $s_i$ とインパルス応答 $h_{ji}$ の畳み込み混合

[数1]

$$x_j(t) = \sum_{i=1}^N \sum_{p=1}^P h_{ji}(p) s_i(t-p+1) \cdots (1)$$

でモデル化される。ここで「畳み込み」とは、信号の伝搬過程で、信号が遅延され、所定の係数が乗算された後、加算されることをいう。また、すべての信号はあるサンプリング周波数でサンプリングされ、離散的に表現されるものとする。そして、式(1)における $P$ はインパルス応答長を、 $t$ はサンプリング時刻を、 $p$ は掃引(時間シフトした信号のサンプル値それぞれに異なる係数を作用させる操作)のための変数を、それぞれ示している。なお、 $N$ 個の信号源701は統計的に互いに独立であり、それぞれの信号は十分スパースであると仮定する。また、「スパース」とは、信号が殆どの時刻 $t$ において0であることを指し、このスパース性は、例えば音声信号で確認される。

[0004] BSSの目的は、源信号 $s_i$ やインパルス応答 $h_{ji}$ を知らずに、観測信号 $x_j$ のみから、分離システム(W)703を推定し分離信号 $y_k$ を得ることである。

また、畳み込み混合の問題は扱いが繁雑であること、さらに、スパース性の仮定は時間-周波数領域でよりよく成立することから、上述の式(1)に短時間離散フーリエ変換(DFT: Discrete Fourier Transform)を施して、信号を時間-周波数領域に変換した上で問題を扱うことが有効である。時間-周波数領域では、上述の式(1)は、

$$X(f, m) = H(f) S(f, m)$$

となる。ここで、 $f$ は周波数、 $m$ はDFTに用いるフレームの時刻を表す。また、 $H(f)$ は、その $ij$ 要素に信号源 $i$ からセンサ $j$ までの周波数応答 $H_{ji}(f)$ を持つ $(M \times N)$ 行列であり、以後これを混合行列と呼ぶ。また、 $S(f, m) = [S_1(f, m), \dots, S_N(f, m)]^T$ 、 $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]^T$ はそれぞれ、源信号と観測信号のDFT結果である。なお、記号 $[\alpha]^T$ は $\alpha$ の転置行列を表す。また、 $S(f, m)$ 及び $X(f, m)$ はベクトルである。

[0005] 以降、時間-周波数領域で説明を行う。

#### <分離過程のモデル>

次に、分離過程についてモデル化する。

まず、 $W(f, m)$ を、その $jk$ 要素にセンサ $j$ での観測信号から分離信号 $y_k$ までの周波数応答 $W_{jk}(f, m)$ を持つ $(N \times M)$ 行列であるとする。この $W(f, m)$ を分離行列と呼ぶ。分離行列を用いると、分離信号は時間-周波数領域で、

$$Y(f, m) = W(f, m) X(f, m)$$



となる。ここで  $Y(f, m) = [Y_1(f, m), \dots, Y_N(f, m)]^T$  は、時間-周波数領域での分離信号であり、これを短時間逆離散フーリエ変換 (IDFT: Inverse Discrete Fourier Transform) することで、源信号の推定結果である分離信号  $y_k$  を得る。なお、分離された分離信号  $y_k$  の順序は、源信号  $s_i$  の順序と必ずしも一致しない。すなわち、 $k=i$  とは限らない。また、 $Y(f, m)$  はベクトルである。

[0006] <分離行列  $W(f, m)$  の推定>

BSSでは、観測信号のみから分離行列  $W(f, m)$  を推定する。

分離信号  $Y(f, m)$  の推定のための従来手法には、(a) 独立成分分析による方法、(b) 信号のスパース性を利用した方法、(c) スパース性により混合行列を推定する方法が知られている。以下、それぞれについて説明を行う。

[従来法1: 独立成分分析による方法]

前述の式(1)のように線形混合された信号を、信号の統計的独立性に基づいて分離する技術は、独立成分分析 (ICA: Independent Component Analysis) と呼ばれる。  $N=M=2$  の場合について、このICAによる分離過程のブロック図を、図27Bに示す。時間-周波数領域のICAでは、出力信号  $Y(f, m)$  の各要素が互いに独立となるよう、学習則  $W(f) = W(f) + \Delta W(f)$  により逐次的に学習を行い、各周波数における分離行列  $W(f, m)$  を求める。ここでは、ICA分離行列推定部705が、例えば、

$$\Delta W = \mu [I - \langle \phi(Y(f, m)) Y(f, m)^H \rangle] \dots (2)$$

という学習則により  $\Delta W(f)$  を求める。但し  $[\alpha]^H$  は  $\alpha$  の共役転置を示す。なお、 $I$  は単位行列、 $\langle \cdot \rangle$  は時間平均、 $\phi$  はある非線形関数、 $\mu$  は更新係数を、それぞれ表す。また、ICAで求められる分離システムは、時不変線形システムとなる。なお、ICAのアルゴリズムは、非特許文献1に記載されているものなど、様々なものが紹介されている。

[0007] ICAでは信号の独立性に着目して分離を行うため、この分離行列  $W(f, m)$  を用い、 $Y'(f, m) = W(f, m)X(f, m)$  によって得られる  $Y'(f, m) = [Y'_1(f, m), \dots, Y'_N(f, m)]^T$  には、順序の任意性と大きさの任意性がある。これは、順序や大きさが変わっても分離信号間の独立性が保たれるからである。

順序の任意性を解くことをパーミュテーション (permutation) の解決と呼ぶが、これ

は、同じ源信号 $s_i$ に対応する分離信号成分が、すべての周波数で同じ添字 $i$ を持つ分離信号 $Y_i(f, m)$ になるようにするものである。その方法としては、分離行列の逆行列( $N \neq M$ の場合はMoore-Penrose型擬似逆行列)を用いて得られる信号の推定到来方向を検証し、 $i$ 番目の分離信号に対応する推定到来方向が各周波数においてすべて同じとなるように分離行列 $W(f, m)$ の行を入れ換える方法や、周波数間で $i$ 番目の分離信号の絶対値 $|Y_i(f, m)|$ の相関が最も高くなるように分離行列 $W(f, m)$ の行を入れ換える方法などがある。なお、この例のパーミュテーション／スケーリング解決部706は、分離信号 $Y_i(f, m)$ をフィードバックしつつ、このパーミュテーションの解決を行う。

- [0008] また、大きさの任意性を解くことをスケーリング(scaling)の解決と呼ぶ。パーミュテーション／スケーリング解決部706は、例えば、permutation解決後に得られている分離行列 $W(f, m)$ の逆行列( $N \neq M$ の場合はMoore-Penrose型擬似逆行列) $W^{-1}(f, m)$ を計算し、分離行列 $W(f, m)$ の各行 $w_i(f, m)$ について

$$w_i(f, m) \leftarrow [W^{-1}(f, m)]_{ji} w_i(f, m)$$

とし、このスケーリングの解決を行う。

そして、順序と大きさの任意性を解決した分離行列 $W(f, m)$ を用い、 $Y(f, m) = W(f, m)X(f, m)$ により各周波数での分離信号を得る。

- [0009] なお、上述の学習則については、例えば、式(2)における非線形関数として、

$$\phi(Y) = \phi(|Y|) \cdot \exp(j \cdot \angle(Y))$$

$$\phi(x) = \text{sign}(x)$$

などを用いることができる。また、上述のように、permutation解決法としては、例えば、信号到来方向推定法や分離信号の周波数の周波数類似度を利用した方法の何れか、若しくは、両者を組み合わせた方法を用いることができ、それについては特許文献1や非特許文献2に詳しい。さらに、ICAでは、信号源の数 $N$ とセンサ数 $M$ が $M \geq N$ の関係にある必要がある。

- [0010] [従来法2:スパース性による方法]

信号源の数 $N$ とセンサ数 $M$ が $M \leq N$ の関係にある場合の分離手法として、信号のスパース性による方法がある(例えば、非特許文献3)。

信号のスパース性と相互独立性を仮定することで、複数の信号が同時に存在していても、サンプルレベルでは、同時刻に互いに重なり合って観測される確率が低いということを仮定できる。すなわち、各時刻における観測信号には、高々1個の信号しか含まれないということを仮定できる。従って、それぞれの時刻で観測された信号が、どの信号源から発せられた信号であるかを何らかの方法で推定し、その時刻の信号のみを抽出するような関数(バイナリマスク)を分離システム $W(f, m)$ として用いることで、信号を分離することが可能である。これがスパース性による方法である。

[0011] 図28(従来法2)は、このスパース性による方法を説明するためのブロック図である。

各時刻での信号源の推定には、以下の方法を用いるのが一般的である。すなわち、それぞれの信号源が空間的に離れて配置されているとすると、複数のセンサで観測される信号間に、各信号源とセンサの相対位置によって決まる位相差や振幅比が発生する。各時刻における観測信号には高々1つの信号しか含まれないという仮定から、各時刻における観測信号の位相差や振幅比は、その時刻の観測信号に含まれる1つの信号の位相や振幅となる。従って、各サンプルにおける観測信号の位相差や振幅比をクラスタリングすることができ、それぞれのクラスに属する時刻の信号を再構成することで各源信号を推定することができる。

[0012] より具体的に述べる。はじめに観測信号相対値計算部751において、観測信号 $X(f, m)$ 間の

[数2]

$$\text{位相差 } z_1(f, m) = \angle \frac{X_i(f, m)}{X_j(f, m)} (i \neq j) \text{ 及び 振幅比 } z_2(f, m) = \frac{|X_i(f, m)|}{|X_j(f, m)|} (i \neq j)$$

の少なくとも一方を計算し、それを相対値 $z(f, m)$ とする。また、或いは位相差そのものではなく、位相差から求められる信号の到来方向を相対値 $z(f, m)$ としてもよい。

[0013] そして、クラスタリング部752で相対値 $z(f, m)$ の分布を調べると、 $N$ 個のクラスを持つ分布となる。図29に、この分布を例示する。なお、この例は、3信号の混合信号( $N=3$ )をセンサ1( $j=1$ )及びセンサ2( $j=2$ )で観測した場合の例であり、図29Aは位相差又は振幅比のみを用いて分布を求めた例、図29Bは位相差と振幅比とを用いて分布を求めた例である。この図に示すように、スパース性により、これらの分布は、

それぞれ $N=3$ 個のクラス801～803或いは811～813に分類できることが分かる。

- [0014] 次に、代表値算出部753において、これら $N$ 個のクラスの代表値(ピーク・平均値・中央値など)を求める。以降記載の便宜上、値の小さい方から $a_1, a_2, \dots, a_N$ とする(図29の場合は $a_1, a_2, a_3$ )。

そして、バイナリマスク作成部754において、

[数3]

$$M_k(f, m) = \begin{cases} 1 & a_k - \varepsilon \leq z(f, m) \leq a_k + \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (k=1, \dots, N) \dots (3)$$

というバイナリマスク $M_k(f, m)$ を作成する。ここで $\varepsilon$ はバイナリマスクの幅を決めるパラメタである。次に、信号抽出部755において $Y_k(f, m) = M_k(f, m)X_j(f, m)$ の演算を行い、 $k$ 番目の分離信号を得る。なお、 $j$ は任意のセンサ番号である。

即ち、この例のスパース性による方法では、分離行列 $W(f, m)$ は時変であり、

$$W_{jk}(f, m) = M_k(f, m) \quad \text{for } j \in \{1, \dots, M\}$$

$$W_{kl}(f, m) = 0 \quad \text{for } l \neq j (l=1, \dots, M)$$

という非線型システムとなる。

- [0015] [従来法3:スパース性により混合行列を推定する方法]

信号源の数 $N$ とセンサ数 $M$ が $M=N$ の関係にある場合の信号分離手法として、信号のスパース性を用いて混合行列 $H(f)$ を推定し、その逆行列を用いて信号を分離する方法がある(例えば、非特許文献4や非特許文献5参照。)

図28(従来法3)は、このスパース性により混合行列を推定する方法を説明するためのブロック図である。

混合信号 $X(f, m)$ は、混合行列 $H(f)$ を用いて

[数4]

$$\begin{bmatrix} X_1(f, m) \\ X_2(f, m) \\ \vdots \\ X_N(f, m) \end{bmatrix} = \begin{bmatrix} H_{11}(f) & \cdots & H_{1N}(f) \\ H_{21}(f) & \cdots & H_{2N}(f) \\ \vdots & \ddots & \vdots \\ H_{N1}(f) & \cdots & H_{NN}(f) \end{bmatrix} \begin{bmatrix} S_1(f, m) \\ S_2(f, m) \\ \vdots \\ S_N(f, m) \end{bmatrix} \cdot \cdot \cdot (4)$$

$$= \begin{bmatrix} 1 & \cdots & 1 \\ H_{21}(f)/H_{11}(f) & \cdots & H_{2N}(f)/H_{1N}(f) \\ \vdots & \ddots & \vdots \\ H_{N1}(f)/H_{11}(f) & \cdots & H_{NN}(f)/H_{1N}(f) \end{bmatrix} \begin{bmatrix} H_{11}(f)S_1(f, m) \\ H_{12}(f)S_2(f, m) \\ \vdots \\ H_{1N}(f)S_N(f, m) \end{bmatrix} \quad (5)$$

$$\equiv \hat{H}(f)\hat{S}(f, m) \cdot \cdot \cdot (6)$$

と表される。そのため、 $\hat{H}(f)$ を推定できれば、

$$Y(f, m) = \hat{S}(f, m) = \hat{H}(f)^{-1} X(f, m) \cdot \cdot \cdot (7)$$

により、分離信号 $Y(f, m)$ を推定できる。以下に、この $\hat{H}(f)$ の推定から分離信号 $Y(f, m)$ を得るまでの流れを説明する。なお、以下において、

$\alpha^\wedge$ は、 $\alpha$ を示す。

[0016] はじめに、[従来法2]と同様な手順により、観測信号相対値計算部751、クラスタリング部752、代表値算出部753、バイナリマスク作成部754及び信号抽出部755において、1つの信号しか存在しない時刻の信号

[数5]

$$\hat{X}(f, m) = M_k(f, m)X(f, m)$$

を得る。ここでは、すべてのセンサの観測信号 $X(f, m) = [X_1(f, m), \cdots, X_M(f, m)]^T$ についてバイナリマスク $M_k(f, m)$ を適用する。このとき例えば、源信号 $S_i(f, m)$ のみがアクティブな時刻 $m_i$ の観測信号は、

[数6]

$$\hat{X}_j(f, m_i) = M_i(f, m_i)X_j(f, m_i) \approx H_{ji}(f)S_i(f, m_i) \cdots (8)$$

となる。

このように求められた分離信号  $\hat{X}_j(f, m_i)$  は混合過程計算部 756 に送られ、そこで

$$\begin{aligned} \hat{H}_{ji}(f) &= E \left[ \frac{M_k(f, m_i)X_j(f, m_i)}{M_k(f, m_i)X_1(f, m_i)} \right] \\ &= E \left[ \frac{\hat{X}_j(f, m_i)}{\hat{X}_1(f, m_i)} \right] = E \left[ \frac{H_{ji}(f)S_i(f, m_i)}{H_{1i}(f)S_i(f, m_i)} \right] = E \left[ \frac{H_{ji}(f)}{H_{1i}(f)} \right] \cdots (9) \end{aligned}$$

を計算することにより  $\hat{H}(f)$  が推定される。ここで、 $E[\cdot]$  は、 $m_i$  に関する平均である。このように求められた  $\hat{H}(f)$  は、逆行列計算部 757 に送られ、そこでその逆行列  $\hat{H}^{-1}(f)$  が求められる。そして、信号分離部 758 において、上述の式 (7) の演算を行うことにより、分離信号  $Y(f, m)$  の推定ができる。

なお、この手法は、 $\hat{H}(f)$  の逆行列を用いるため、信号源の数  $N$  とセンサ数  $M$  が  $M = N$  の関係にある場合にしか適用できない。

特許文献1: 特開2004-145172号公報

非特許文献1: A. Hyvaerinen and J. Karhunen and E. Oja, "Independent Component Analysis," John Wiley & Sons, 2001, ISBN 0-471-40540

非特許文献2: H. Sawada, R. Mukai, S. Araki and S. Makino, "A Robust and Precise Method for Solving the Permutation Problem of Frequency-Domain Blind Source Separation", in Proc. the 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003), 2003, pp. 505-510

非特許文献3: S. Rickard, R. Balan, and J. Rosca, "Real-Time Time-Frequency Based Blind Source Separation," 3rd International Conference on Independent Component Analysis and Blind Source Separation (ICA2001), San Diego, December, 2001, pp. 651-656

非特許文献4: F. Abrard, Y. Deville, P. White, "From blind source separation to blind source cancellation in the underdetermined case: a new approach based on

time-frequency analysis," Proceedings of the 3rd International Conference on Independent Component Analysis and Signal Separation (ICA'2001), pp. 734-739, San Diego, California, Dec. 2001.

非特許文献5: Y. Deville, "Temporal and time-frequency correlation-based blind source separation methods," in Proc. , ICASSP2003, Apr. 2003, pp. 1059-1064

## 発明の開示

### 発明が解決しようとする課題

[0017] 従来の信号分離方法では、信号源の数 $N$ とセンサの数 $M$ が $N > M$ の関係にある場合に、混合信号を高い品質で分離することは困難であった。

つまり、前述のように、信号源の数 $N$ とセンサの数 $M$ が $N > M$ の関係にある場合、独立成分分析による方法、及びスパース性により混合行列を推定する方法は使用できない。

また、信号のスパース性を利用した方法は使用できるが、この方法では、分離性能がよく、なおかつ歪みが小さい信号分離を行うことが困難である。つまり、上述の式(3)で示されるバイナリマスクの作成時、 $\varepsilon$ を十分小さくすると良い分離性能を得ることができるが、その反面、このバイナリマスクによって排除されるサンプルの数が増加し、分離信号が劣化する。すなわち、信号のスパース性が完全なのであれば、各時刻の観測信号には高々1個の信号しか含まれず、各時刻における各相対値 $z(f, m)$ は、何れかの代表値 $a_1, \dots, a_N$ の近傍に収まるはずである。しかし、実際には信号のスパース性は完全ではないため、ある時刻において、同一周波数の観測信号が2個以上存在する場合もある。この場合、この時刻における相対値 $z(f, m)$ は、本来対応すべき代表値 $a_1, \dots, a_N$ から離れた値となり、 $\varepsilon$ の値によってはバイナリマスクによって排除されてしまう。その結果、このサンプルに対応する観測信号が0として取り扱われ、分離信号に0成分が詰め込まれることになる。そして、この排除されるサンプルの割合は $\varepsilon$ の値が小さいほど大きいため、この0成分が詰め込まれる量も $\varepsilon$ の値が小さいほど大きくなる。そして、各分離信号に多くの0成分が詰め込まれた場合、これが原因となって、分離信号の歪みが大きくなり、Musical Noiseと呼ばれる聴感上不快なノイズが発生する。一方、バイナリマスクの $\varepsilon$ を大きくすると、分離信号に詰められる0成

分が少なくなりMusical Noiseは減少するが、その反面、分離性能が劣化する。

- [0018] 本発明はこのような点に鑑みてなされたものであり、信号源の数 $N$ とセンサの数 $M$ が $N > M$ の関係にある場合でも、混合信号を高い品質で分離することが可能な技術を提供することを目的とする。

#### 課題を解決するための手段

- [0019] 第1の本発明では、以下のように上記課題を解決する。

まず、 $M$ 個のセンサで観測された $N$  ( $N \geq 2$ ) 個の信号の混合である観測信号の値を周波数領域値に変換し、その周波数領域値を用い、センサ間における観測値の相対値(相対値の写像も含む)を、各周波数において算出する。そして、これらの相対値を $N$ 個のクラスにクラスタリングし、それらの各クラスの代表値を算出する。その後、それらの代表値を用い、周波数領域値から $V$  ( $V \leq M$ ) 個の信号源から発せられた信号の値を抽出するためのマスクを作成し、生成したマスクを用い、当該 $V$ 個の信号源から発せられた信号から成る限定信号の値を抽出する。そして、 $V \geq 2$ の場合には、この限定信号は $V$ 個の信号源から発せられた信号から成る混合信号となるため、この限定信号をさらに分離して各分離信号の値を得る。一方、 $V = 1$ の場合には、この限定信号の値を分離信号の値とする。

- [0020] ここで抽出された $V$ 個の信号源から発せられた信号からなる限定信号の分離には、例えば独立成分分析による方法やスパース性により混合行列を推定する方法等を適用できる。そのため $N > M$ の場合でも高い品質で源信号を抽出することができる。ただし、これだけでは $V$ 個の源信号しか抽出できない。そこで、例えば、複数種類のマスクを用い、抽出する信号の組合せを変化させながら同様な処理を繰り返すことにより、すべての源信号を抽出する。

また、第2の本発明では、以下のように上記課題を解決する。

- [0021] まず観測信号値 $x_1(t), \dots, x_M(t)$ を周波数領域値 $X_1(f, m), \dots, X_M(f, m)$ に変換する。そして、周波数領域値 $X_1(f, m), \dots, X_M(f, m)$ からなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$ を、周波数 $f$ ごとに $N$ 個ずつのクラス $C_i(f)$  ( $i = 1, \dots, N$ )にクラスタリングし、各クラス $C_i(f)$ を代表する第2のベクトル $a_i(f)$ を算出し、そこから $V$  ( $V \leq M$ ) 個の第3のベクトル $a_p(f)$  ( $p = 1, \dots, V$ )を抽出する



。その後、第3のベクトル $a_p(f)$ の集合を $G_k$ とし、 $G_k^c$ を $G_k$ の補集合とし、 $D(\alpha, \beta)$ をベクトル $\alpha$ と $\beta$ とのマハラノビス平方距離とした場合における、

[数7]

$$M(f, m) = \begin{cases} 1 & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

で示されるマスク $M(f, m)$ を生成し、マスク $M(f, m)$ と第1のベクトル $X(f, m)$ との積を演算して、 $V$ 個の信号源から発せられた信号からなる限定信号の値を抽出する。

[0022] ここで抽出された $V$ 個の信号源から発せられた信号からなる限定信号の分離には、例えば独立成分分析による方法やスパース性により混合行列を推定する方法等を適用できる。そのため $N > M$ の場合でも高い品質で源信号を抽出することができる。ただし、これだけでは $V$ 個の源信号しか抽出できない。そこで、例えば、複数種類の集合 $G_k$ に対する複数種類のマスクを用い、抽出する信号の組合せを変化させながら同様な処理を繰り返す。これにより、すべての源信号を抽出する。

[0023] また、第3の本発明では、以下のように上記課題を解決する。

まず、観測信号値 $x_1(t), \dots, x_M(t)$ を、周波数領域値 $X_1(f, m), \dots, X_M(f, m)$ に変換し、それらからなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]^T$ を、周波数 $f$ ごとに $N$ 個ずつのクラスタ $C_i(f)$  ( $i=1, \dots, N$ )にクラスタリングする。なお、源信号がスパースであれば、たとえセンサの数が不十分( $N > M$ )の状況でも、 $N$ 個のクラスタ $C_i(f)$ にクラスタリングでき、それら $N$ 個の代表ベクトル $a_i(f)$ を算出することも可能である。

[0024] そして、それらの各クラスタ $C_i(f)$ を代表する第2のベクトル $a_i(f)$ を算出し、 $N$ 個の第2のベクトル $a_i(f)$ の中の0個以上の当該第2のベクトルを0ベクトルに置換した $M$ 行 $N$ 列の行列 $A'$ のムーア・ペンローズ(Moore-Penrose)型擬似逆行列( $A'^+ (f) : N=M$ の場合は逆行列 $A'^{-1} (f)$ に一致)である $N$ 行 $M$ 列の分離行列 $W(f, m)$ を算出する。なお、ここで生成される分離行列 $W(f, m)$ は、センサの数が不十分( $N > M$ )である場合には時間 $m$ に依存する行列となり、センサの数が十分( $N \leq M$ )である場合には時間 $m$ に依存しない行列となる。

[0025] その後、 $Y(f, m) = W(f, m)X(f, m)$ の演算により、分離信号ベクトル $Y(f, m) = [$

$Y_1(f, m), \dots, Y_N(f, m)]^T$ を算出し、時間領域の信号値 $y_1(t), \dots, y_N(t)$ に変換する。

ここで、源信号のスパース性により、たとえ信号源の数 $N$ がセンサの数 $M$ よりも多い( $N > M$ ) 場合であっても、離散時間 $m$ ごとに見れば、観測結果に影響を及ぼす値をとる信号源の数は $M$ 個以下である可能性が高い。そのため、離散時間 $m$ ごとに見れば、上述のように生成された分離行列 $W(f, m)$ によって、これら $M$ 個以下の信号を分離することは可能である。そして、 $N > M$ である場合、分離行列 $W(f, m)$ は時間依存となるのだから、離散時間 $m$ が相違すれば、得られる分離信号の組合せも相違しうる。そのため、複数の離散時間 $m$ について分離信号を求めていくことにより、すべての分離信号を得ることもできる。

### 発明の効果

- [0026] 以上のように、本発明では、信号源の数 $N$ とセンサの数 $M$ が $N > M$ の関係にある場合でも、混合信号を高い品質で分離することができる。

### 図面の簡単な説明

- [0027] [図1]第1の実施の形態における信号分離装置の全体構成を例示したブロック図。  
[図2]図1における代表値生成部、マスク制御部、限定信号作成部及び限定信号分離部の詳細を例示したブロック図。  
[図3]図1及び図2のマスク作成部の詳細を例示したブロック図。  
[図4]第1の実施の形態における信号分離装置の処理を説明するためのフローチャート。  
[図5]クラスタリング部により作成されたヒストグラムの例示。  
[図6]第1の実施の形態における滑らかな形状のマスクを生成する際使用する信号の推定到来方向 $\theta_i$ のとり方を説明するための図。  
[図7]第1の実施の形態におけるマスクの例示。  
[図8]第2の実施の形態の信号分離装置の1系統を例示したブロック図。  
[図9]第3の実施の形態の信号分離装置の1系統を例示したブロック図。  
[図10]第3の実施の形態におけるマスクの例示。  
[図11]第4の実施の形態におけるマスク作成部の構成を例示したブロック図。

[図12]Aは、第6の実施の形態におけるバイナリマスクの例示、Bは、第7の実施の形態におけるバイナリマスクの例示。

[図13]第8の実施の形態における代表値生成部、マスク制御部及び限定信号作成部の構成を例示したブロック図。

[図14]第8の実施の形態における信号分離処理を説明するためのフローチャート。

[図15]第9の実施の形態の信号分離装置の構成を例示したブロック図。

[図16]第9の実施の形態の信号分離装置の処理を説明するためのフローチャート。

[図17]センサの数が不十分な場合( $M < N$ )における分離行列生成処理を説明するためのフローチャート。

[図18]1音源の場合の正規化していない観測信号ベクトル $X(f, m)$ のプロット。

[図19]1音源の場合において、式(36)により正規化した観測信号ベクトル $X(f, m)$ のプロット。

[図20]1音源の場合において、式(37)により正規化した観測信号ベクトル $X(f, m)$ のプロット。

[図21]2音源の場合の正規化していない観測信号ベクトル $X(f, m)$ のプロット。

[図22]2音源の場合において、式(36)により正規化した観測信号ベクトル $X(f, m)$ のプロット。

[図23]2音源の場合において、式(37)により正規化した観測信号ベクトル $X(f, m)$ のプロット。

[図24]センサの数が信号源の数に対して十分であるか否かに係わらず適用できる分離行列生成処理を説明するためのフローチャート。

[図25]周波数領域で信号統合を行ってから時間領域に変換する際の構成を例示したブロック図の一部。

[図26]各実施の形態をコンピュータで構成した信号分離装置の例。

[図27]Aは、従来のブラインド信号分離技術を概念的に例示したブロック図、Bは、ICAによる分離過程のブロック図。

[図28]スパース性による方法及びスパース性により混合行列を推定する方法を説明するためのブロック図。

[図29]相対値の分布の例示。

### 符号の説明

- [0028] 1, 500 信号分離装置  
2, 501 記憶部  
3, 502 信号分離プロセッサ

### 発明を実施するための最良の形態

- [0029] 以下、本発明の実施の形態を図面を参照して説明する。

#### 〔第1の実施の形態〕

本形態は、第1の本発明に係る実施の形態であり、死角型ビームフォーマの指向特性を用いた滑らかな形状のマスクを用い、観測信号値から、 $V(2 \leq V \leq M)$ 個の信号源から発せられた信号から成る混合信号(本形態ではこれを「限定信号」と呼ぶ)の値を抽出し、抽出した限定信号値をICAによって信号分離する例である。

図1は、本形態の信号分離装置1の全体構成を例示したブロック図である。また、図2は、図1における代表値生成部30、マスク制御部40、限定信号作成部50-k( $k=1, \dots, u$ , 但し $u$ は後述する系統数)及び限定信号分離部60-kの詳細を例示したブロック図である。また、図3は、図1及び図2のマスク作成部51-kの詳細を例示したブロック図である。なお、これらの図における矢印はデータの流れを示すが、制御部10や一時記憶部90に出入りするデータの流れは省略してある。すなわち、データが制御部10や一時記憶部90を経由する場合であっても、その経由の過程は省略してある。また、図4は、本形態における信号分離装置1の処理を説明するためのフローチャートである。以下、これらの図を用いて、この例の信号分離装置1の構成及び処理を説明していく。

- [0030] <全体構成>

まず、本形態の信号分離装置の全体構成について説明する。

図1に例示するように、本形態の信号分離装置1は、記憶部2とこれに有線或いは無線で電氣的に接続された信号分離プロセッサ3とを有している。

記憶部2は、例えば、ハードディスク装置、フレキシブルディスク、磁気テープ等の磁気記録装置、DVD-RAM(Random Access Memory)、CD-R(Recordable)/R

W (ReWritable)等の光ディスク装置、MO (Magneto-Optical disc)等の光磁気記録装置、EEP-ROM (Electrically Erasable and Programmable-Read Only Memory)、フラッシュメモリ (flash memory)等の半導体メモリ等である。また、記憶部2は、信号分離プロセッサ3と同一の筐体内に存在してもよいし、別個の筐体に構成されてもよい。

[0031] またこの例の信号分離プロセッサ3は、例えば、プロセッサやRAM等によって構成されるハードウェアであり、以下に述べる各処理ブロックを有する。

＜信号分離処理の概略＞

次に、信号分離装置1が行う信号分離処理の概略について説明する。

本形態では、N個の信号源から発せられた信号は統計的に互いに独立であり、それぞれの信号は十分スパースであると仮定する。ここで「スパース」とは、信号が殆どの時刻 $t$ において0又は0に近く、大きな値をとることは稀であるという性質である。このスパース性は、例えば音声信号で確認される。なお、音声信号など白色でない信号は、短時間離散フーリエ変換等を施して周波数ごとの時間系列とすることで、より0に近い時刻が増えてスパース性が強調される。また、一般には信号のモデル化にガウス分布が用いられることが多いが、スパース性を持つ信号はガウス分布ではなくラプラス分布などでモデル化される。

[0032] まずM個の観測信号値 $x_j(t)$ を周波数領域変換部20にて周波数領域の観測信号値 $X_j(f, m)$ に変換した後、代表値生成部30において、各源信号に対応するN個の代表値 $a_1, a_2, \dots, a_N$ を算出する。

次に、マスク制御部40にて代表値 $a_1, a_2, \dots, a_N$ のうち $V$  ( $2 \leq V \leq M$ ) 個を適当に選び、限定信号作成部50-kにおいて、観測信号値 $X_j(f, m)$ からV個の源信号のみから構成される限定信号の値 $\hat{X}(f, m)$ を推定する。なお、 $V=1$ の場合には、後述する[第3の実施の形態]の方法を用いる。ここでは、マスク作成部51-kにおいてV個の信号を取りだすような滑らかな形状のマスクを作成し、限定信号抽出部52-kでこのマスクを観測信号値 $X_j(f, m)$ に作用させることで、限定信号値 $\hat{X}(f, m)$ を推定する。

[0033] 次に限定信号分離部60-kにおいて、V個の分離信号を得るための分離システムを

推定する。ここでは、 $M$ 個の限定信号値 $\hat{X}(f, m)$ を入力とし、 $V$ 個の分離信号値 $Y(f, m)$ を出力する。ここで、分離システムの入力数 $M$ と出力数 $V$ について、 $V \leq M$ であるので、ここでの分離システムの推定には、[従来法1]や[従来法3]を用いることが可能である。

[0034] 最後に時間領域変換部70-kにおいて、時間周波数領域で得られている分離信号値 $Y(f, m)$ を時間領域の信号値に変換する。

しかし以上の処理だけでは $V$ 個の分離信号しか得られない。よって、その他の分離信号を得るために、マスク制御部40で選択する $V$ 個の代表値の構成を変え、限定信号作成部50-kから時間領域変換部70-kまでの処理を複数系統( $u$ 系統)行う。

そして最後に、信号統合部80にて、各系統からの出力を統合し、 $N$ 個全ての分離信号を得る。

[0035] <構成及び処理の詳細>

次に、この例の構成及び処理の詳細について説明する。

この例は、 $N$  ( $N \geq 2$ ) 個の信号源から発せられた信号が混合し、 $M$ 個のセンサにおいて観測される状況において、その観測信号から源信号を分離抽出するものである。なお、上述のようにこの例の信号は音声信号等のスパース性を仮定できる信号であり、その音源数 $N$ は既知或いは推定可能とする。また、この例のセンサは、この信号を観測できるマイクロホン等であり、それらは直線上に配置されるものとする。

[0036] まず、前処理として、各センサにおいて観測された時間領域の各観測信号 $x_j(t)$  ( $j = 1, \dots, M$ )を記憶部2に格納しておく。そして、信号分離処理が開始されると、信号分離プロセッサ3は、制御部10の制御のもと以下の処理を実行する。

まず信号分離プロセッサ3は、記憶部2にアクセスし、そこから各観測信号値 $x_j(t)$ を順次読み込み、周波数領域変換部20に送る(ステップS1)。周波数領域変換部20は、短時間離散フーリエ変換等によって、これらの信号値を時間ごとの周波数領域の観測信号値 $X_j(f, m)$ に順次変換し、一時記憶部90に格納する(ステップS2)。一時記憶部90に格納された周波数領域の観測信号値 $X_j(f, m)$ は、代表値生成部30に送られ、代表値生成部30の相対値算出部31は、送られた周波数領域の観測信号値 $X_j(f, m)$ を用い、各センサ間における観測値の相対値 $z(f, m)$ を、各周波数にお

いて算出する(ステップS3)。

[0037] なお、相対値 $z(f, m)$ としては、例えば、

[数8]

$$\text{位相差 } z_1(f, m) = \angle \frac{X_i(f, m)}{X_j(f, m)} (i \neq j) \text{ 及び振幅比 } z_2(f, m) = \frac{|X_i(f, m)|}{|X_j(f, m)|} (i \neq j)$$

の少なくとも一方を用いてもよく、また、或いは位相差そのものではなく、その写像(例えば、位相差から求められる信号の到来方向)を用いてもよい。

この例では、j1番目とj2番目の任意の2つのセンサにおける観測信号間位相差 $z_1(f, m)$ から得られる信号の到来方向

[数9]

$$z_3(f, m) = \cos^{-1} \frac{z_1(f, m)v}{2\pi fd}$$

をこの相対値 $z(f, m)$ として利用し、相対値算出部31は、この $z_3(f, m)$ を算出するものとする。ここで $v$ は信号の早さ、 $d$ はセンサj1とセンサj2との間隔である。

このように算出された相対値 $z_3(f, m)$ は、一時記憶部90に格納される。次に、クラスタリング部32は、一時記憶部90から相対値 $z_3(f, m)$ を順次読み込み、これら相対値 $z_3(f, m)$ をN個のクラスにクラスタリングする(ステップS4)。この例の場合、クラスタリング部32は、送られた相対値 $z_3(f, m)$ からヒストグラムを作成する。

[0038] 図5は、このように作成されたヒストグラムの例示である。なお、この例は源信号の数 $N=3$ のものである。

この図に例示するように、この例のヒストグラムは、 $N(=3)$ 個のピークを持つ分布になる。この例のクラスタリング部32は、この分布を $N(=3)$ 個のクラス(この例ではクラス91〜93)にクラスタリングする。これは、例えば、適当な閾値を設定してクラスタリングしてもよいし、k-means法や階層的クラスタリング等の多くの教科書で説明されている方法を用いてもよい(例えば、「尾上守夫 監訳“パターン識別”，新技術コミュニケーションズ，ISBN 4-915851-24-9，第10章」等参照)。ここでクラスタリングされた各クラス $C_i (i=1, 2, \dots, N)$ は、相対値 $z_3(f, m)$ の集合であり、離散時間の集合 $T_i$ を用いて $C_i(f) = \{z_3(f, m) \mid m \in T_i\}$ である。

[0039] クラスタリング部32で生成されたクラスタリングの情報(クラス $C_1, C_2, \dots, C_N$ )は、一時記憶部90に格納される。代表値計算部33は、これらを読み込み、N個の各クラス $C_1, C_2, \dots, C_N$ の代表値 $a_1, a_2, \dots, a_N$ を算出する(ステップS5)。具体的には、例えば、ヒストグラムの各クラスのピークを代表値としてもよく、また、各クラスの平均値を代表値としてもよい。そして、例えばこのN個の代表値を、(便宜上)値の小さい方から $a_1, a_2, \dots, a_N$ とする(図5参照)。なお、これらの代表値 $a_1, a_2, \dots, a_N$ は、N個の各信号の到来方向の推定値になっている。

[0040] この例の代表値 $a_1, a_2, \dots, a_N$ の情報は、一時記憶部90に格納された後、マスク制御部40に送られる。マスク制御部40は、これら代表値 $a_1, a_2, \dots, a_N$ を要素に持つ集合 $G_0$ を特定するデータを変数 $SG_0$ に代入し、この変数 $SG_0$ を一時記憶部90に格納する。また、マスク制御部40は、集合Gを特定する変数SGを $G = \phi$  (空集合)に初期化し、変数kを0とし、それらを一時記憶部90に格納する(ステップS6)。

次に、マスク制御部40での制御のもと、N個すべての分離信号が得られるまで、限定信号作成部50-k ( $k=1, \dots, u$ )、限定信号分離部60-k及び時間領域変換部70-kの複数系統(u系統)による処理が行われる。

[0041] まず、マスク制御部40は、一時記憶部90に格納された変数kに1を加えた値を新たな変数kとし、再び一時記憶部90に格納する(ステップS7)。次に、マスク制御部40は、一時記憶部90から変数 $SG_0$ 及びSGを呼び出す。そして、マスク制御部40は、変数 $SG_0$ によって特定される集合 $G_0$ から、SGによって特定される集合Gの補集合( $G^c$  ( $\alpha^c$ は $\alpha$ の補集合を示す))の元を含む適当な $V (\leq M)$ 個の代表値の集合 $G_k$ を選択し、この集合 $G_k$ を特定するデータを変数 $SG_k$ に代入し、この変数 $SG_k$ を一時記憶部90に格納する(ステップS8)。

[0042] 限定信号作成部50-kのマスク作成部51-kは、一時記憶部90に格納された変数 $SG_k$ を読み出し、この変数 $SG_k$ によって特定される集合 $G_k$ を代表値に持つクラスの信号を抽出する「滑らかな形状のマスク」を作成する(ステップS9)。ここで、「滑らかな形状のマスク」とは、 $V (2 \leq V \leq M)$ 個の代表値を含む所定の範囲(限定範囲)内にある相対値に対してハイレベル値をとり、この限定範囲内にはない代表値に対してローレベル値をとり、相対値の変化に伴う当該ハイレベル値から当該ローレベル値への推移



が連続的である関数を意味する。なお、この例の「ハイレベル値」とは、0より十分大きな数値(例えば1以上)を意味し、「ローレベル値」とは0に十分近い値(例えば、ハイレベル値に対して60dB以下等)を意味するが、特にその値に限定はない。

- [0043] 本形態では、 $N-V+1$ 個のセンサにより形成される死角型ビームフォーマの指向特性を利用して「滑らかな形状のマスク」を作成する。このマスクは、限定信号に含まれる $V$ 個の信号の方向( $G_k$ )へは十分な感度を持ち、除去されるべき $N-V$ 個の信号の方向( $G_0 \cap G_k^c$ )へは感度の低い特性(死角)を持つ、なめらかな形状のマスクである。

以下に、本形態の「滑らかな形状のマスク」の生成手順について説明する。

まず、マスク作成部51-kが、一時記憶部90から変数 $SG_k$ 、 $SG_0$  及び  $SG_k^c$  を読み出す。次に、マスク作成部51-kは、変数 $SG_k$  が示す集合 $G_k$  の要素(限定範囲内の代表値)の何れか一つを抽出し、これを  $\theta_1$  とする。また、マスク作成部51-kは、変数 $SG_0$  及び  $SG_k^c$  によって特定される $G_0 \cap G_k^c$  の要素(限定範囲内にない代表値)すべてを抽出し、これらを  $\theta_i$  ( $i=2, \dots, N-V+1$ ) とする。そして、マスク作成部51-kは、 $\theta_1$  及び  $\theta_i$  を一時記憶部90に格納する。次に、マスク作成部51-kは、一時記憶部90から  $\theta_1$  及び  $\theta_i$  を順次抽出し、 $\tau_{ji} = (d_j/v) \cos \theta_i$  ( $j=1, \dots, N-V+1$ ) を順次算出する。さらに、マスク作成部51-kは、遅延行列  $H_{NBF}(f)$  の  $ji$  要素  $H_{NBFji}(f) = \exp(j2\pi f \tau_{ji})$  を算出して一時記憶部90に順次格納する。なお、 $d_j$  はセンサ1とセンサ $j$ との距離( $d_1$  は0)であり、 $f$  は周波数の変数、 $v$  は信号の速さである。これらのパラメータは、例えば事前に一時記憶部90に格納され、順次呼び出されて使用される。以上の処理により、 $((N-V+1) \times (N-V+1))$  の遅延行列  $H_{NBF}(f)$  が生成される(図3:51a-k)。

- [0044] なお、本形態では、2つのセンサにおける観測信号間位相差 $z_1(f, m)$  から得られる信号の到来方向 $z_3(f, m)$  を相対値としているため、上述の  $\theta_1$  は限定範囲内の代表値に対応する信号の到来方向を示し、 $\theta_i$  は、限定範囲内にない代表値に対応する信号の到来方向を示している。また、この  $\theta_i$  ( $i=1, 2, \dots, N-V+1$ ) のとり方は図6のようになる。まず、直線上に配置された $M$ 個のセンサの中央を原点とする(1番目のセンサと原点との距離 $L_1$  = 原点と $M$ 番目のセンサとの距離 $L_2$ )。この原点と $i$ 番目の

信号源とを結ぶ線分と、原点と1番目のセンサ10とを結ぶ線分とがなす角度が、 $i$ 番目の信号源に対応する  $\theta_i$  である。

- [0045] 生成された遅延行列  $H_{\text{NBF}}(f)$  は、一時記憶部90(図1)からNBF作成部51b-k(図3)に送られ、NBF作成部51b-kは、この遅延行列  $H_{\text{NBF}}(f)$  を用い、死角ビームフォーマ(NBF)の特性を持つNBF行列  $W(f)$  を作成する。これは、遅延行列  $H_{\text{NBF}}(f)$  の逆行列  $W(f) = H_{\text{NBF}}^{-1}(f)$  を算出することで得られる。

このNBF行列  $W(f)$  は、一時記憶部90(図1)に格納される。指向特性計算部51c-kは、このNBF行列  $W(f)$  の1行目の要素  $W_{1k}(f)$ 、 $d_k$  及び  $v$  を一時記憶部90から順次抽出し、 $\theta$  を信号の到来方向の変数とした場合における、指向特性関数

[数10]

$$F(f, \theta) = \sum_{k=1}^{N-V+1} W_{1k}(f) \exp(j2\pi f d_k \cos \theta / v) \quad \cdots (10)$$

を生成する。なお、 $\theta$  の取り方は上述の  $\theta_i$  と同じである。

- [0046] 生成された指向特性関数  $F(f, \theta)$  は、マスク構成部51d-kに送られる。マスク構成部51d-kは、この指向特性関数  $F(f, \theta)$  と、一時記憶部90から読み出した相対値  $z(f, m)$  (この例では  $z_3(f, m)$ ) とを用い、滑らかな形状のマスク  $M_{\text{DC}}(f, m)$  を生成する。

- [0047] 生成するマスク  $M_{\text{DC}}(f, m)$  としては、例えば、この指向特性  $F(f, \theta)$  自身

$$[\text{マスク1}] \quad M_{\text{DC}}(f, m) = F(f, z_3(f, m)) \quad \cdots (11)$$

を用いたものを例示できる。

また、指向特性  $F(f, \theta)$  の絶対値

$$[\text{マスク2}] \quad M_{\text{DC}}(f, m) = |F(f, z_3(f, m))| \quad \cdots (12)$$

を用いたものをマスク  $M_{\text{DC}}(f, m)$  としてもよい。

図7Aは、[マスク2]の例(信号数  $N=3$ 、センサ数  $M=2$  の場合)を示している。この例の「滑らかな形状のマスク」は、信号除去数  $N-M=1$  のものであり、一方向  $a_1$  へ小さいゲインを持つものである。なお、この「滑らかな形状のマスク」は、 $M(=V)=2$  個の信号(ここでは  $a_2$  及び  $a_3$  方向より到来する2個の信号)を限定信号として抽出するためのものである(後述の図7B, Cも同様)。

[0048] また、例えば以下のように、指向特性 $F(f, \theta)$ を変形したものをマスク $M_{DC}(f, m)$ としてもよい。なお以下では、 $G_k$ の要素の中の互いに隣り合う2つの $a_i$ に挟まれる相対値 $z_3(f, m)$ の領域すべてを限定信号領域と呼ぶ。また、 $G_k$ に $a_1$ や $a_N$ が含まれる場合、 $0^\circ \leq z_3(f, m) \leq a_1$ 、 $180^\circ \geq z_3(f, m) \geq a_N$ も限定信号領域に含むものとする。さらに、 $G_0 \cap G_k^c$ の要素の中の互いに隣り合う2つの $a_i$ に挟まれる相対値 $z_3(f, m)$ の領域すべてを除去信号領域と呼ぶ。また、 $G_0 \cap G_k^c$ に $a_1$ や $a_N$ が含まれる場合、 $0^\circ \leq z_3(f, m) \leq a_1$ 、 $180^\circ \geq z_3(f, m) \geq a_N$ も除去信号領域に含むものとする。そして、限定信号領域、除去信号領域のどちらにも属さない領域を過渡領域と呼ぶ。

[0049] [数11]

[マスク 3]

$$M_{DC}(f, m) = \begin{cases} F(f, z_3(f, m)) & z_3(f, m) \in \text{除去信号領域以外の領域} \\ F(f, \theta_r) & z_3(f, m) \in \text{除去信号領域} \end{cases} \quad \dots (13)$$

[マスク 4]

$$M_{DC}(f, m) = \begin{cases} |F(f, z_3(f, m))| & z_3(f, m) \in \text{除去信号領域以外の領域} \\ |F(f, \theta_r)| & z_3(f, m) \in \text{除去信号領域} \end{cases} \quad \dots$$

(14)

これらは、除去信号領域のゲインを一様に小さくする特性を持つマスクをマスク $M_{DC}(f, m)$ としたものである。ここで $\theta_r$ は、除去信号領域の端点のうち、隣り合う限定信号領域の端点と最も近いものを意味する。図7Bに、この[マスク4]の例(信号数 $N=3$ 、センサ数 $M=2$ の場合)を示す。

[0050] また、例えば、

[数12]

[マスク 5]

$$M_{DC}(f, m) = \begin{cases} a & z_3(f, m) \in \text{限定信号領域} \\ b & z_3(f, m) \in \text{除去信号領域} \cdots (15) \\ F(f, z_3(f, m)) & z_3(f, m) \in \text{過渡領域} \end{cases}$$

のように、限定信号領域の指向特性を一様にしたマスク  $M_{DC}(f, m)$  を用いることも可能である。また、

[マスク 6]

$$M_{DC}(f, m) = \begin{cases} a & z_3(f, m) \in \text{限定信号領域} \\ b & z_3(f, m) \in \text{除去信号領域} \cdots (16) \\ |F(f, z_3(f, m))| & z_3(f, m) \in \text{過渡領域} \end{cases}$$

のように、限定信号領域の指向特性を一様にしたマスクの絶対値を用いることもできる。

ここでaとしては例えば、除去信号領域の  $|F(f, \theta)|$  の最大値などの0より十分大きい値を用い、bとしては例えば、指向特性のゲインの最小値などの小さな値を用いる。図7Cに[マスク6]の例(信号数 $N=3$ 、センサ数 $M=2$ の場合)を示す(マスク作成部51-k/ステップS9の説明終わり)。

[0051] 以上のようにマスク作成部51-kで生成されたマスク  $M_{DC}(f, m)$  は、一時記憶部90に格納された後、限定信号抽出部52-kに送られる。限定信号抽出部52-kは、さらに一時記憶部90から周波数領域の観測信号値  $X(f, m)$  を読み出す。そして、限定信号抽出部52-k(図2)は、このマスク  $M_{DC}(f, m)$  と周波数領域の観測信号値  $X(f, m)$  とを用い、 $X_k^{\wedge}(f, m) = M_{DC}(f, m) X(f, m)$  の積演算により、限定信号値  $X_k^{\wedge}(f, m)$  を生成する(ステップS10)。

[0052] この限定信号値  $X_k^{\wedge}(f, m)$  は、一時記憶部90に格納され、限定信号分離部60-kは、この限定信号値  $X_k^{\wedge}(f, m)$  を読み出し、限定信号の信号分離を行う(ステップS11)。ここで、限定信号値  $X_k^{\wedge}(f, m) = M_{DC}(f, m) X(f, m)$  は、 $V(2 \leq V \leq M)$  個の信号源から発せられた信号によって構成された混合信号の値であると近似される。よって、その分離行列の推定には[従来法1]で述べた独立成分分析による方法を利用できる。すなわち独立成分分析の入力として、観測信号値  $X$  の代わりに限定信号値  $X_k^{\wedge}(f, m)$  を用い、例えば[従来法1]で述べた式(2)を用いて分離を行う。

[0053] 本実施例におけるICAによる分離では、まず、ICA分離行列推定部61-kにおいて、限定信号値 $\hat{X}_k(f, m)$ を用い、前述の式(2)の学習則に従い分離行列 $W(f, m)$ を生成し、この分離行列 $W(f, m)$ を一時記憶部90に格納する。なお、この分離行列 $W(f, m)$ の生成には、例えば、以下のパーミュテーション・スケーリング解決部62-kからの出力値 $Y_k(f, m)$ のフィードバックを用いる。生成された分離行列 $W(f, m)$ はパーミュテーション・スケーリング解決部62-kに送られる。パーミュテーション・スケーリング解決部62-kは、この分離行列 $W(f, m)$ と限定信号値 $\hat{X}_k(f, m)$ を用い、 $Y_k(f, m) = W(f, m) \hat{X}_k(f, m)$ の演算を行い、それぞれの分離信号値 $Y_k(f, m) = [Y_{k1}^{\Pi k1}(f, m), \dots, Y_{kV}^{\Pi kV}(f, m)]^T$ を生成し、それを一時記憶部90に格納する。そして、パーミュテーション・スケーリング解決部62-kは、例えば、この分離信号値 $Y_k(f, m)$ をフィードバックし、[従来法1]で述べた方法でPermutation問題を解決する。Permutation問題の解決後、パーミュテーション・スケーリング解決部62-kは、さらに、分離信号値 $Y_{kq}(q=1, \dots, V)$ がどの源信号に対応するのかを示すタグ $\Pi_{kq}$ を、分離信号値 $Y_{kq}(q=1, \dots, V)$ に付与し、これらに対応付けて一時記憶部90に格納する。ここでは、このタグ $\Pi_{kq}$ を分離信号値 $Y_{kq}$ の上付添字 $\Pi_{kq}$ として表記する。

[0054] 具体的には、例えば、パーミュテーション・スケーリング解決部62-kが、一時記憶部90から抽出した分離行列 $W(f)$ の逆行列( $N \neq M$ の場合はMoore-Penrose型擬似逆行列)を用い、

[数13]

$$\theta_q = \arccos \frac{\arg([W_{jq}^{-1}]/[W_{jq}^{-1}])}{2\pi f v^{-1} d} \dots (17)$$

[0055] (但し $v$ は信号の速さ、 $d$ はセンサ $j$ とセンサ $j'$ との間隔)

の演算によって得られる信号の推定到来方向 $\theta_q$ と、一時記憶部90から抽出した変数 $SG_k$ が示す集合 $G_k$ に含まれる代表値とを比較し、 $\theta_q$ に最も近い代表値 $a_i$ を $q$ 番目の分離信号 $Y_{kq}$ に対応付ける(ステップS12)。つまり、パーミュテーション・スケーリング解決部62-kは、この分離信号 $Y_{kq}$ に対し、代表値 $a_i$ を示すタグ $\Pi_{kq}$ を付与する(対応付ける)。

[0056] また、この後、パーミュテーション・スケーリング解決部62-kが、一時記憶部90から

分離行列 $W(f)$ を抽出し、その各行 $w_q(f)$ を

$$w_q(f) \leftarrow [W^{-1}(f)]_{jq} w_q(f)$$

とすることにより、ICAのスケーリング問題を解決し、スケーリング問題解決後の分離行列 $W(f)$ を一時記憶部90に格納する。なお、後の信号統合部80における処理のため、この処理ではすべての系列 $k$ において同じ $j$ を用いることが望ましい。

- [0057] タグ $\Pi_{kq}$ が付与された各分離信号値 $Y_{kq}$ は、時間領域変換部70-kに送られる。時間領域変換部70-kは、例えば、短時間逆離散フーリエ変換等により、時間周波数領域で得られている各分離信号値 $Y_{kq}$ を時間領域の信号値に変換し、その変換値を一時記憶部90に格納する。(ステップS13)。なお、これら時間領域の信号値 $y_k(t) = [y_{k1}^{\Pi_{k1}}(t), \dots, y_{kV}^{\Pi_{kV}}(t)]^T$ にも上述のタグ $\Pi_{kq}$ が関連付けられる。この関連付けを行う場合、まず、時間領域変換部70-kが、一時記憶部90から、周波数領域の信号値 $Y_{kq}$ に対応付けられているタグ $\Pi_{kq}$ を各周波数について抽出する。次に、時間領域変換部70-kは、各周波数におけるタグ $\Pi_{kq}$ がすべて等しいか否かを判断する。ここでこれらがすべて等しかった場合には、時間領域の信号値 $y_{kq}$ のタグとして、周波数領域の信号値 $Y_{kq}$ に対応付けられているタグ $\Pi_{kq}$ を対応付ける。一方、これらがすべて等しくなかった場合には、多数決にて時間領域の信号値 $y_{kq}$ のタグを決定する。

- [0058] 次に、マスク制御部40において、一時記憶部90から変数 $SG$ と $SG_k$ とを抽出し、これらが示す $G$ と $G_k$ との和集合 $G \cup G_k$ を新たな集合 $G$ とし、この集合 $G$ を変数 $SG$ に代入し、この変数 $SG$ を一時記憶部90に格納する(ステップS14)。また、マスク制御部40は、一時記憶部90から変数 $SG$ と $SG_0$ とを読み出し、この新たな集合 $G$ が集合 $G_0$ と等しいか否かを判断する(ステップS15)。ここで、 $G = G_0$ でなければステップS7の処理に戻る。

- [0059] 一方、 $G = G_0$ であれば、信号統合部80において、一時記憶部90から各系統 $k$ (時間領域変換部70-k/ $k=1, \dots, u$ )から出力された分離信号 $y_{kp}(t)$ を読み出し、これらの選択/統合を行い、 $N$ 個すべての分離信号を得る(ステップS16)。具体的には、例えば、まず信号統合部80は、一時記憶部90から読み出した各分離信号 $y_{kp}(t)$ のタグ $\Pi_{kq}$ を比較する。ここで、複数の系統 $k$ において同じタグを持つ分離信号値 $y_{kp}(t)$ が存在しないと判断された場合、信号統合部80は、すべての分離信号値 $y_{kp}(t)$

を最終的な分離信号値 $y_i(t)$  ( $i=1, \dots, N$ )として出力する(ステップS17)。一方、複数の系統において同じタグを持つ分離信号値が存在すると判断された場合、信号統合部80は、これらのタグが等しい分離信号値のどれか1つを適当に選択し、最終的な分離信号値 $y_i(t)$ として出力するか、同じタグを持つ分離信号値の平均を計算し、それを出力信号とする(ステップS17)。

[0060] ここで、どれか1つの分離信号値 $y_{kq}(t)$ を適当に選択し、最終的な分離信号値 $y_i(t)$ として出力する処理の場合、信号統合部80は、例えば、同じタグ $a_i$ を持つ分離信号値 $y_{kq}(t)$ の中で最大パワーを持つものを最終的な分離信号値 $y_i(t)$ として出力する。また、同じタグを持つ分離信号値の平均を最終的な分離信号値 $y_i(t)$ として出力する処理の場合、信号統合部80は、例えば、

[0061] [数14]

$$y_i(t) = \frac{1}{K} \sum_{\Pi_{kq=a_i}} y_{kq}^{\Pi_{kq}}(t)$$

( $K$ は同じタグ $a_i$ を持つ分離信号の個数)

とする。以上により、 $N$ 個の信号が少ない歪で分離される。

[0062] <本形態の特徴>

従来の「従来法2:信号のスパース性を利用した方法」で分離性能を上げた際に分離信号の歪が大きくなるのは、分離性能を上げるために上述の式(3)の $\varepsilon$ を十分小さくすると、このバイナリマスクによって取り出される信号成分が制限され、本来原信号の成分として取り出されるべきサンプルの多くが取り出されないからである。つまり、この場合、各分離信号に多くの0成分が詰めこまれ、各分離信号を不連続とし、Musical Noiseを発生させることになる。

[0063] これに対し、この形態では、滑らかな形状を持つマスクによって、2個以上 $M$ 個以下の原信号からなる混合信号(限定信号)を抽出する。そのため、1個のみの信号の値を抽出する[従来法2]のバイナリマスクよりも、広い範囲の相対値 $z(f, m)$ に対する信号(サンプル)を限定信号と抽出できる。

そのため、ある時刻において同一周波数の観測信号が2個以上存在し、サンプル値が、本来対応すべき代表値から離れてしまった場合であっても、このようなサン

ル値を抽出できる可能性は高くなる。その結果、分離信号に不連続に0成分が詰めこまれることによる品質劣化 (Musical Noiseの発生) を抑制できる。

[0064] また、本形態では $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測される状況において、滑らかな形状を持つマスクを用いて信号の分離抽出を行った。この滑らかな形状を持つマスクは、[従来法2]によるマスク(0或いは1の値をとるバイナリマスク)と異なり、そのエッジ部分が滑らかに広がった形状を有する。そのため、この滑らかな形状のマスクを用いれば、ある時刻において同一周波数の観測信号が2個以上存在し、サンプル値が、本来対応すべき代表値 $a_1, \dots, a_N$ から離れた場合であっても、この位置に対するマスクが0以外の値を持つ場合もあるため、急峻に値が変化するバイナリマスクよりも、多くの信号を抽出することができる。その結果、分離信号に不連続に0成分が詰めこまれることによる品質劣化を抑制できる。

[0065] また、滑らかな形状のマスクはエッジ部分に近づくほど値が小さくなるため、従来のバイナリマスクにおいて単に $\epsilon$ を大きくした場合に比べ、分離性能の劣化も少ない。

さらに、抽出される限定信号は、 $V$  ( $\leq M$ ) 個の源信号のみからなると考えられるため、分離問題が簡単になっている。従って、限定信号については、[従来法1]や[従来法3]を用いて容易に信号の分離ができる。また、後述の第3の実施の形態で説明するように、 $V=1$ の場合には、[従来法1]や[従来法3]の方法を用いる必要もない。

[0066] <性能比較>

以下は、[従来法2]で信号分離を行った場合と、本形態の方法で[マスク2]を用いて信号分離を行った場合との性能を比較した表である。

[表1]

	SIR1	SIR2	SIR3	SDR1	SDR2	SDR3
従来法2	17.3	11.6	17.6	8.1	7.4	7.1
本形態	-	5.9	17.6	-	13.4	17.4
	18.5	7.0	-	16.2	13.0	-

この例では、源信号として、3人の話者(男性2名・女性1名)による音声信号を用い、残響の無い環境でのこれらの混合信号を、2つの無指向性マイクで観測する状況をシミュレートしている。なお、表中のSIRは信号対妨害音比 (Signal to interference ratio) (dB) であり、分離性能を示す指標である。また、SDRは信号対歪比 (Signal to



distortion ratio) (dB) であり、信号の歪の程度を示す指標である。双方とも値が大きい方が性能が良いことを示している。また、SIR1及びSDR1は話者1に、SIR2及びSDR2は話者2に、SIR3及びSDR3は話者3に、それぞれ対応している。また、本形態のデータは縦2段になっているが、これは $k=1$ の系統の分離結果と、 $k=2$ の系統の分離結果にそれぞれ対応するものである。

- [0067] この表に示すように、この形態の方法では、分離性能SIRをほとんど落すことなく、従来法2に比べ格段に高いSDRを得ることができている。これは信号の歪が少ない分離ができていることを示している。これより、本形態の方法は、信号源の数 $N$ がセンサの数 $M$ より多い場合に信号を低歪で分離するために有効であることが分かる。

〔第2の実施の形態〕

本形態も第1の本発明に係る実施の形態である。本形態では、限定信号作成部で「滑らかな形状のマスク」を用い、限定信号分離部で混合行列推定による分離方法を用いる例を示す。なお、本形態において第1の実施の形態と共通する事項については説明を省略する。

- [0068] 図8は、本形態の信号分離装置のうち、 $V$ 個の分離信号値を得る1系統のみを例示したブロック図である。

なお、図8において第1の実施の形態と共通する構成については、第1の実施の形態と同じ符号を付した。図8に例示するように、第1の実施の形態の信号分離装置1と本形態の信号分離装置と構成上の相違点は、限定信号作成部50- $k$ が限定信号作成部150- $k$ に置き換わり、限定信号分離部60- $k$ が限定信号分離部160- $k$ に置き換わる点である。また、マスク作成部151- $k$ が2種類のマスクを作成する点、 $V=M$ に限定される点も異なる。以下、本形態の構成及び処理について説明する。

- [0069] まず、代表値生成部30(図8)は、周波数領域変換部20(図1)が生成した周波数領域の観測信号値 $X_j(f, m)$ を一時記憶部90から抽出する。次に、代表値生成部30(図8)は、第1の実施の形態と同様に、相対値算出部31において観測値の相対値 $z(f, m)$ を算出し、クラスタリング部32においてクラスタリングを行い、代表値算出部33において代表値 $a_1, a_2, \dots, a_N$ を算出する。なお、本形態では、相対値 $z(f, m)$ として、 $i$ 番目と $j$ 番目の任意の2つのセンサにおける観測信号間位相差 $z_{ij}(f, m)$ から得られる

信号の到来方向

[数15]

$$z_3(f, m) = \cos^{-1} \frac{z_1(f, m)v}{2\pi fd}$$

を用いるのが望ましい。

- [0070] これらの代表値 $a_1, a_2, \dots, a_N$ は、一時記憶部90(図1)に格納された後、マスク制御部40(図8)を介し、限定信号作成部150-kのマスク作成部151-kに送られ、マスク作成部151-kは、2種類のマスクを作成する。1つは、 $G_k$ に含まれる $V(=M)$ 個の代表値に対応する $V(=M)$ 個の信号が混合した限定信号の値 $\hat{X}(f, m)$ を抽出するためのマスクであり、第1の実施の形態に示した滑らかな形状のマスク $M_{DC}(f, m)$ である。もう1つは、1個の信号のみを含む信号を抽出するバイナリマスク $M_k(f, m)$ であり、[従来法2]で示したものと同様のマスク

[数16]

$$M_k(f, m) = \begin{cases} 1 & a_k - \epsilon \leq z(f, m) \leq a_k + \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (k=1, \dots, N) \dots (18)$$

である。これらのマスクは一時記憶部90(図1)に格納される。

- [0071] 次に、限定信号抽出部152-k(図8)は、一時記憶部90(図1)から滑らかな形状のマスク $M_{DC}(f, m)$ と周波数領域の観測信号値 $X(f, m)$ とを読み出す。そして、限定信号抽出部152-k(図8)は、このマスク $M_{DC}(f, m)$ を周波数領域の観測信号値 $X(f, m)$ に掛けた限定信号値 $\hat{X}(f, m) = M_{DC}(f, m)X(f, m)$ を算出し、これを一時記憶部90(図1)に格納する。ここで、この限定信号値 $\hat{X}(f, m)$ は、 $V$ 個の信号が混合したものと近似されるので、限定信号分離部160-kにおける信号の分離には[従来法3]で述べた混合行列推定法が応用できる。

- [0072] そこでまず、限定信号分離部160-kの積演算部161-k(図8)において、一時記憶部90(図1)からバイナリマスク $M_k(f, m)$ と周波数領域の観測信号値 $X(f, m)$ とを読み出す。そして、積演算部161-k(図8)は、 $\hat{X}_k(f, m) = M_k(f, m)X(f, m)$ の演算を行い、1個の信号のみを含む分離信号の値 $\hat{X}_k(f, m)$ を求め、これを一時記憶部90(図1)に格納する。次に、混合過程推定部162-k(図8)が、一時記憶部90(図

1)から $\hat{X}_k(f, m)$ を読み出し、[従来法3]と同様に

[数17]

$$\begin{aligned}\hat{H}_{ji}(f) &= E \left[ \frac{M_i(f, m_i) X_j(f, m_i)}{M_i(f, m_i) X_i(f, m_i)} \right] \\ &= E \left[ \frac{\hat{X}_j(f, m_i)}{\hat{X}_i(f, m_i)} \right] = E \left[ \frac{H_{ji}(f) S_i(f, m_i)}{H_{ii}(f) S_i(f, m_i)} \right] = E \left[ \frac{H_{ji}(f)}{H_{ii}(f)} \right]\end{aligned}$$

によって推定された混合行列 $\hat{H}$ を算出する。なお、この混合行列 $\hat{H}$ のサイズは $N \times M$ となっている。なお、この混合行列は全ての系列 $k$ で求める必要はなく、ある系列で推定した $\hat{H}$ を一時記憶部90に格納し、逐次読み出して用いてもよい。

[0073] この混合行列 $\hat{H}$ は逆行列計算部163-kに送られ、逆行列計算部163-kは、まずこの混合行列 $\hat{H}$ ランクを落す。すなわち、混合行列 $\hat{H}$ のうち、 $V$ 個の信号からなる限定信号 $\hat{X}(f, m)$ に対応する $V$ 列(すなわち、 $G_k$ に含まれる $V$ 個の代表値 $a_i$ に対応する列)のみを取り出し、 $V \times V$ の正方行列 $\hat{H}_M$ を作成する。これは、 $V$ 個の信号の混合と近似される限定信号 $\hat{X}(f, m)$ の分離を行うためである。

[0074] 次に逆行列計算部163-kは、作成した正方行列 $\hat{H}_M$ の逆行列 $\hat{H}_M^{-1}(f)$ を計算し、これを一時記憶部90(図1)に格納する。積演算部164-k(図8)は、限定信号値 $\hat{X}(f, m)$ と逆行列 $\hat{H}_M^{-1}(f)$ とを一時記憶部90(図1)から読み出し、 $Y_k(f, m) = \hat{H}_M^{-1}(f) \hat{X}(f, m)$ の演算により、 $V$ 個の分離信号値の推定値 $Y_k(f, m) = [Y_{k1}^{\Pi k1}(f, m), \dots, Y_{kv}^{\Pi kv}(f, m)]^T$ を算出する。なお、分離信号 $Y_{kq}$  ( $q=1, \dots, V$ )がどの源信号に対応するのかを示すタグ情報の付与は、前述の式(17)において、 $W^{-1}$ の代わりに $\hat{H}_M$ を用いて信号の推定到来方向を求め、その方向がどの代表値 $a_i$ に近いかを判断して行う。

[0075] [第3の実施の形態]

本形態も第1の本発明に係る実施の形態である。本形態では、「滑らかな形状のマスク」を用い、観測信号から、何れか1個の信号源から発せられた信号によって構成される信号(本形態ではこれを「限定信号」と呼ぶ)のみを抽出し、抽出した限定信号を分離信号とする。なお、本形態において第1の実施の形態と共通する事項につい

ては説明を省略する。

図9は、本形態の信号分離装置のうち、1個の分離信号を得る1系統部分のみを例示したブロック図である。なお、図9において第1の実施の形態と共通する構成については、第1の実施の形態と同じ符号を付した。

[0076] 図9に例示するように、第1の実施の形態の信号分離装置1と、本形態の信号分離装置と構成上の相違点は、限定信号作成部50-kが限定信号作成部250-kに置き換わる点、及び本形態の信号分離装置には限定信号分離部60-kが存在しない点である。以下、本形態の構成及び処理について説明する。

まず、代表値生成部30(図9)は、周波数領域変換部20が生成した周波数領域の観測信号値 $X_j(f, m)$ を一時記憶部90(図1)から抽出する。代表値生成部30(図9)は、第1の実施の形態と同様、相対値算出部31において観測値の相対値 $z(f, m)$ を算出し、クラスタリング部32においてクラスタリングを行い、代表値算出部33において代表値 $a_1, a_2, \dots, a_N$ を算出する。なお、相対値 $z(f, m)$ としては、位相差及び振幅比の少なくとも一方、或いはその写像(例えば、位相差から求められる信号の到来方向)等を用いることができるが、本形態では、観測信号間位相差から求められる信号の到来方向

[数18]

$$z_3(f, m) = \cos^{-1} \frac{z_1(f, m)v}{2\pi fd}$$

を相対値とする。

[0077] これらの代表値 $a_1, a_2, \dots, a_N$ は、一時記憶部90(図1)に格納され、限定信号作成部250-kのマスク作成部251-k(図9)は、これらの代表値 $a_1, a_2, \dots, a_N$ を読み出し、何れかの一つの代表値 $a_i$ を抽出するための「滑らかな形状のマスク」を生成する。なお、本形態の「滑らかな形状のマスク」は、 $V(V=1)$ 個の代表値を含む限定範囲内にある相対値に対してハイレベル値をとり、この限定範囲内にない代表値に対してローレベル値をとり、相対値の変化に伴う当該ハイレベル値から当該ローレベル値への推移が連続的である関数である。

[0078] 以下に、本形態の「滑らかな形状のマスク」の生成手順について説明する。

まず、マスク作成部251-kは、 $(N \times N)$ の遅延行列 $H_{NBF}(f)$ を生成する。すなわち、マスク作成部251-kは、一時記憶部90(図1)に格納された代表値 $a_1, a_2, \dots, a_N$ のうちの一つ(抽出する信号の到来方向の推定値)を抽出し、これを $\theta_1$ とする。また、マスク作成部251-kは、それ以外の $N-1$ 個の代表値(抽出しない信号の到来方向の推定値)を一時記憶部90(図1)から抽出し、それらを $\theta_i$  ( $i=2, \dots, N$ )とする。これらの $\theta_1$ 及び $\theta_i$ は一時記憶部90(図1)に格納される。マスク作成部251-kは、一時記憶部90から $\theta_1$ 及び $\theta_i$ を順次抽出し、 $\tau_{ji} = (d_j/v) \cos \theta_i$  ( $j=1, \dots, N$ )を算出し、遅延行列 $H_{NBF}(f)$ の $ji$ 要素 $H_{NBFji}(f) = \exp(j2\pi f \tau_{ji})$ を算出して一時記憶部90に順次格納する。なお、 $d_j$ はセンサ1とセンサ $j$ との距離( $d_1$ は0)であり、 $f$ は周波数の変数、 $v$ は信号の速さである。これらのパラメータは、例えば事前に一時記憶部90に格納され、順次呼び出されて使用される。以上の処理により、 $(N \times N)$ の遅延行列 $H_{NBF}(f)$ が生成される。

[0079] 次に、マスク作成部251-kは、この遅延行列 $H_{NBF}(f)$ を用い、死角ビームフォーマ(NBF)の特性を持つNBF行列 $W(f)$ を作成する。これは、遅延行列 $H_{NBF}(f)$ の逆行列 $W(f) = H_{NBF}^{-1}(f)$ を算出することで得られる。この逆行列 $W(f) = H_{NBF}^{-1}(f)$ は一時記憶部90に格納される。そして、マスク作成部251-kは、一時記憶部90からこのNBF行列 $W(f)$ の1行目の要素 $W_{1k}(f)$ 、 $d_k$ 及び $v$ を一時記憶部90から順次抽出し、前述の式(10)に示した指向特性関数 $F(f, \theta)$ を生成する。その後、マスク作成部251-kは、この指向特性関数 $F(f, \theta)$ を用い、滑らかな形状のマスク $M_{DC}(f, m)$ を生成する。

[0080] 具体的には、例えば、第1の実施の形態における式(11)で示されるマスク([マスク7]とする)や、式(12)で示されるマスク([マスク8]とする)を、本形態の滑らかな形状のマスク $M_{DC}(f, m)$ として生成する。

また、例えば、以下のように除去信号領域のゲインを一様に小さくする特性を持つ[滑らかな形状のマスク]を生成することとしてもよい。

[数19]

[マスク 9]

$$M_{DC}(f, m) = \begin{cases} F(f, z_3(f, m)) & z_3(f, m) \in \text{除去信号以外の領域} \\ F(f, \theta_r) & z_3(f, m) \in \text{除去信号領域} \end{cases} \quad \dots(19)$$

[マスク 10]

$$M_{DC}(f, m) = \begin{cases} |F(f, z_3(f, m))| & z_3(f, m) \in \text{除去信号以外の領域} \\ |F(f, \theta_r)| & z_3(f, m) \in \text{除去信号の領域} \end{cases} \quad \dots(20)$$

[0081]  $\theta_r$  は、除去するN-1個の信号の到来方向の推定値(抽出する代表値 $a_i$ 以外のN-1個の代表値)のうち、除去しない信号の到来方向の推定値(抽出する代表値 $a_i$ )に一番近いものである。

また、例えば、

[数20]

[マスク 11]

$$M_{DC}(f, m) = \begin{cases} |F(f, \theta_1)| & z_3(f, m) \in \text{除去信号以外の領域} \\ |F(f, \theta_r)| & z_3(f, m) \in \text{除去信号領域} \\ F(f, z_3(f, m)) & z_3(f, m) \in \text{過渡領域} \end{cases} \quad \dots(21)$$

のように、取り出す方向についての指向特性を均一にしたマスク $M_{DC}(f, m)$ を用いることも可能である。また、過渡領域では $M_{DC}(f, m) = |F(f, z_3(f, m))|$ を用いることもできる([マスク12])。

[0082] 図10に、上述の[マスク8][マスク12]の例を示す。これらは、信号数 $N=3$ 、センサ数 $M=2$ の場合に、 $a_1$ の方向から到来する信号を抽出し、 $a_2$ 及び $a_3$ の方向から到来する信号を抑圧する「滑らかな形状のマスク」の例である。

マスク生成部251-kで生成された滑らかな形状のマスク $M_{DC}(f, m)$ は、限定信号抽出部252-kに送られ、限定信号抽出部252-kは、 $Y_k(f, m) = M_{DC}(f, m) X_j(f, m)$ により分離信号 $Y_k(f, m)$ を抽出する。以上の処理はすべての分離信号が抽出されるまで複数の系統で実施され、最終的にすべての分離信号 $Y(f, m)$ が得られる。そして、信号分離装置は、得られた分離信号 $Y(f, m)$ を、時間領域変換部において時間領域の信号に戻し、信号統合部をそのまま通過して出力する。

[0083] <性能比較>

以下は、[従来法2]で信号分離を行った場合と、本形態の方法で[マスク8][マスク

11]を用いて信号分離を行った場合との性能を比較した表である。

[表2]

	SIR1	SDR1	SIR2	SDR2	SIR3	SDR3
従来法 2	15.0	7.9	10.3	11.1	17.3	9.0
本形態(マスク 8)	14.8	12.1	5.9	17.2	14.6	11.1
本形態(マスク 11)	15.4	13.0	8.3	16.1	16.1	11.4

この例では、源信号として、3人の話者(男性2名・女性1名)による音声信号を用い、残響の無い環境でのこれらの混合信号を、2つの無指向性マイクで観測する状況をシミュレートしている。

[表3]

	SIR1	SDR1	SIR2	SDR2	SIR3	SDR3
従来法 2	15.1	11.3	9.0	13.3	13.4	9.2
本形態(マスク 8)	14.6	11.4	5.5	17.2	14.3	11.6
本形態(マスク 11)	15.5	12.2	7.9	16.0	15.4	11.7

この例は、表2の状況において信号の混合の仕方(具体的には話者の位置配置)を変えた場合のシミュレート結果である。

[表4]

	SIR1	SDR1	SIR2	SDR2	SIR3	SDR3
従来法 2	11.0	7.7	4.3	10.8	13.4	6.4
本形態(マスク 8)	10.8	7.8	2.7	16.5	12.9	7.6
本形態(マスク 11)	12.0	8.7	3.5	15.7	14.9	7.1

この例は、表2の状況において話者の組合せ(男性3名)を変えた場合の例である。

[0084] これらの表に示すように、この形態の方法では、分離性能SIRをほとんど落すことなく、従来法2に比べ格段に高いSDRを得ることができている。これは信号の歪が少ない分離ができていることを示している。これより、本形態の方法は、信号源の数Nがセンサの数Mより多い場合に信号を低歪で分離するために有効であることが分かる。

[0085] [第4の実施の形態]

本形態も第1の本発明に係る実施の形態である。本形態では、バイナリマスクに滑らかな形状の関数を畳み込んで、滑らかな形状のマスクを生成する。以下ではマスク作成部(図1におけるマスク生成部51-kに相当)における処理のみを説明する。な

お、その他の構成や処理については、第1から第3の実施の形態と同様である。また、本形態では、第1の実施の形態で説明した位相差 $z_1(f, m)$ 、振幅比 $z_2(f, m)$ 、位相差 $z_3(f, m)$ から得られる信号の到来方向 $z_3(f, m)$ 等を、相対値 $z(f, m)$ として使用することができる。

- [0086] 図11は、本形態におけるマスク作成部300-kの構成を例示したブロック図である。マスク作成部300-kの処理が開始されると、まず、バイナリマスク作成部301-kは、V個の代表値を含む所定の範囲内にある相対値に対してハイレベル値をとり、この範囲内にない相対値に対してローレベル値をとり、相対値の変化に伴う当該ハイレベル値から当該ローレベル値への推移が不連続な関数であるバイナリマスクを生成する。例えば、マスク作成部300-kは、V個の信号が混合した信号を抽出するためのバイナリマスク

[数21]

$$F_b(z) = \begin{cases} 1 & a_{\min} \leq z \leq a_{\max} \\ 0 & \text{otherwise} \end{cases}$$

を生成する。

- [0087] なお、 $a_{k+1}$  から  $a_{k+V}$  のV個の代表値を含む信号を抽出する場合、 $a_{\min}$ 、 $a_{\max}$  は、例えば  $a_k < a_{\min} < a_{k+1}$ 、 $a_{k+V} < a_{\max} < a_{k+V+1}$  の範囲で設定する。これらは適当に設定しても良いが、より具体的には、例えば、以下の処理によって  $a_{\min}$ 、 $a_{\max}$  を算出する。

- [0088] まず、マスク作成部300-kは、一時記憶部90(図1)に格納されている相対値 $z(f, m)$ 、クラスタ $C_i$ 及び代表値 $a_i$  ( $i=1, \dots, N$ ) (第1の実施の形態ステップS3-5参照)を読み込み、各クラスタ $C_i$ の分散値を

[数22]

$$\sigma^2(f)_i = (1/|C_i|) \sum_{m \in T_i} (z(f, m) - a_i(f))^2 \quad \dots (22)$$

の演算によって算出する。なお、 $|C_i|$ とは、クラスタ $C_i$ に属する相対値 $z(f, m)$ の数である。また、この分散値の算出を、例えば、EMアルゴリズム(例えば、「尾上守夫 監訳「パターン識別」, 新技術コミュニケーションズ, ISBN 4-915851-24-9, 第10章」等参照。)などを用い、データにガウシアンモデルのあてはめを行って求めても



よい。

算出された分散値  $\sigma_i^2$  は一時記憶部90(図1)に格納され、次に、マスク作成部301-k(図11)は、一時記憶部90に格納されている分散値  $\sigma_i^2$  及び代表値  $a_i$  (この例ではクラス  $C_i$  の平均値)を読み込み、これらを用いて、

[数23]

$$\begin{aligned} a_{\min} &= \frac{\sigma_{k+1} \cdot a_k + \sigma_k \cdot a_{k+1}}{\sigma_{k+1} + \sigma_k} \\ a_{\max} &= \frac{\sigma_{k+v} \cdot a_{k+v+1} + \sigma_{k+v+1} \cdot a_{k+v}}{\sigma_{k+v} + \sigma_{k+v+1}} \end{aligned} \quad \dots(23)$$

を算出する( $a_{\min}$ 、 $a_{\max}$  の具体的な算出例の説明終わり)。

[0089] 以上のように生成されたバイナリマスク  $F_b(z)$  は一時記憶部90(図1)に格納される。

次に、単峰性関数生成部302-k(図11)が、 $z$  の変化に伴って値が連続的に変化する単峰性関数  $g(z)$  を生成し、一時記憶部90(図1)に格納する。なお、単峰性関数  $g(z)$  としては、例えば、ガウシアン

[数24]

$$g(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left\{\frac{z}{\sigma}\right\}^2\right\}$$

等の滑らかな形状の関数を例示できる。なお、 $\sigma$  は  $g(z)$  の標準偏差を意味する。例えば、 $a_{k+1} - \sigma > a_k + \sigma$ 、 $a_{k+v} + \sigma < a_{k+v+1} - \sigma$  となるように適当に設定するのが望ましく、例えば、 $\sigma = \min(\sigma_k, \sigma_{k+v+1})$  とできる。なお、 $\sigma_k$  及び  $\sigma_{k+v+1}$  は式(22)のものである。また、 $\min(\alpha, \beta)$  は  $\alpha$  及び  $\beta$  のうち小さい方を取り出す操作を意味する。

[0090] 次に、畳み込み混合部303-k(図11)が、一時記憶部90(図1)から、バイナリマスク  $F_b(z)$  及び単峰性関数  $g(z)$  を読み込み、このバイナリマスク  $F_b(z)$  に単峰性関数  $g(z)$  を畳み込んだ関数  $F(z) = F_b(z) * g(z)$  を計算し、これを一時記憶部90(図1)に格納する。ここで  $*$  は  $z$  に関する畳み込み演算子である。

そして、次に、マスク構成部304-k(図11)が、一時記憶部90(図1)から、相対値  $z(f, m)$  及び関数  $F(z)$  を読み込み、関数  $F(z)$  に相対値  $z(f, m)$  を代入したマスク

$$M_{DC}(f, m) = F(z(f, m)) \cdots (24)$$

を生成し、一時記憶部90(図1)に格納する。

[0091] なお、その他、バイナリマスク $F_b(z)$ の両端に、ある傾きを直線(曲線)を付加した形状のマスクを滑らかな形状の関数を $F(z)$ とし、式(24)のマスクを求めてもよい。

[0092] また、代表値 $a_i$ (この例ではクラス $C_i$ の平均値)と、式(22)(23)のように求めた分散 $\sigma_i^2$ 及び $a_{\min}$ 、 $a_{\max}$ とをマスク構成部304-k(図11)が読み込み、平均 $a_i(f)$ 、分散 $\sigma_i^2(f)$ をもつガウシアン

[数25]

$$g_i(z) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{1}{2}\left\{\frac{z-a_i}{\sigma_i}\right\}^2\right\}$$

を算出し、さらに $g_i(z)$ を $g_i(z) / g_i(a_i)$ にて、 $a_i$ における値を1に正規化し、

$$F(z) = \begin{cases} g_k(z) & a_{\min} > z \\ 1 & a_{\min} \leq z \leq a_{\max} \\ g_{k+v} & a_{\max} < z \end{cases}$$

を算出し、式(24)のマスクを求めてもよい。

[0093] [第5の実施の形態]

本形態も第1の本発明に係る実施の形態である。本形態は、奇関数の差から滑らかな形状のマスクを生成する。以下ではマスク作成部(図1におけるマスク生成部51-kに相当)における処理のみを説明する。なお、その他の構成や処理については、第1から第3の実施の形態と同様である。

本形態のマスク作成部は、相対値が限定範囲の下限值 $a_{\min}$ である場合に0をとる第1の奇関数と、相対値が限定範囲の上限値 $a_{\max}$ である場合に0をとる第2の奇関数との差の写像から得られる単峰性の関数を、滑らかな形状のマスクとして生成する。例えば、

$M_{DC}(f, m) = \{\tanh(z(f, m) - a_{\min}) - \tanh(z(f, m) - a_{\max})\}^\alpha$ を「滑らかな形状のマスク」とする。なお、相対値 $z(f, m)$ には、第1の実施の形態等で示した位相差 $z_1(f, m)$ 及び振幅比 $z_2(f, m)$ の少なくとも一方、或いはその写像(例えば、位相差から

求められる信号の到来方向 $z_3(f, m)$ 等を用いる。また、 $\alpha$ は任意の正の数であり、 $a_{\min}$ 、 $a_{\max}$ は第4の実施の形態と同様に求められたものである。また、必要に応じて、

$$M_{DC}(f, m) = M_{DC}(f, m) / \max(M_{DC}(f, m))$$

などの正規化を施してもよい。

[0094] 〔第6の実施の形態〕

本形態も第1の本発明に係る実施の形態である。本形態のマスクは、図1及び図2のマスク作成部51-kにおいて作成され、V個の代表値を含む所定の範囲内にある相対値に対してハイレベル値をとり、この所定の範囲内にない代表値に対してローレベル値をとり、ハイレベル値からローレベル値への推移が不連続な関数(バイナリマスク)である。ただし、 $2 \leq V \leq M$ である。すなわち、例えば、

[数26]

$$B(f, m) = \begin{cases} 1 & a_{\min} \leq z(f, m) \leq a_{\max} \\ 0 & \text{otherwise} \end{cases} \quad \dots (25)$$

をバイナリマスクとして作成する。なお、 $a_{k+1}$ から $a_{k+V}$ のV個の代表値を含む信号を抽出する場合、 $a_{\min}$ 、 $a_{\max}$ は、例えば $a_k < a_{\min} < a_{k+1}$ 、 $a_{k+V} < a_{\max} < a_{k+V+1}$ の範囲で設定する。より具体的には、例えば第4の実施の形態で述べた方法と同様な手順により、 $a_{\min}$ 、 $a_{\max}$ を生成する。また、本形態でも、位相差 $z_1(f, m)$ 、振幅比 $z_2(f, m)$ 、位相差 $z_1(f, m)$ から得られる信号の到来方向 $z_3(f, m)$ 等を、相対値 $z(f, m)$ として使用することができる。

[0095] また、 $a_{\min}$ から $a_{\max}$ の範囲に含まれる相対値 $z(f, m)$ の数は、センサの数2以上M以下であり、好ましくは、センサの数Mである。さらに、第1の実施の形態と同様、本形態では複数通りのバイナリマスク $B(f, m)$ を作成する。

[0096] 具体的には、例えば、マスク制御部40(図1, 図2)が一時記憶部90から代表値 $a_1, a_2, \dots, a_N$ を読み出し、これら代表値 $a_1, a_2, \dots, a_N$ を要素に持つ集合 $G_0$ を特定するデータを変数 $SG_0$ に代入し、この変数 $SG_0$ を一時記憶部90に格納する。また、マスク制御部40は、集合 $G$ を特定する変数 $SG$ を $G = \phi$ (空集合)に初期化し、変数 $k$ を0とし、それらを一時記憶部90に格納する(図4:ステップS6)。次に、マスク制御部40での制御のもと、N個すべての分離信号が得られるまで、限定信号作成部50-k( $k=1$ ,

..., u), 限定信号分離部60-k及び時間領域変換部70-kの複数系統(u系統)による処理が行われる。まず、マスク制御部40は、一時記憶部90に格納された変数kに1を加えた値を新たな変数kとし、再び一時記憶部90に格納する(図4:ステップS7)。次に、マスク制御部40は、一時記憶部90から変数 $SG_0$ 及び $SG$ を呼び出す。そして、マスク制御部40は、変数 $SG_0$ によって特定される集合 $G_0$ から、 $SG$ によって特定される集合 $G$ の補集合( $G^c$ ( $\alpha^c$ は $\alpha$ の補集合を示す))の元を含む適当な $V$ ( $\leq M$ )個の代表値の集合 $G_k$ を選択し、この集合 $G_k$ を特定するデータを変数 $SG_k$ に代入し、この変数 $SG_k$ を一時記憶部90に格納する(図4:ステップS8)。限定信号作成部50-kのマスク作成部51-kは、一時記憶部90に格納された変数 $SG_k$ を読み出し、この変数 $SG_k$ によって特定される集合 $G_k$ を代表値に持つクラスの信号を抽出するバイナリマスクを作成する(図4:ステップS9)。

[0097] 図12Aは、本形態におけるバイナリマスクの例示である。この例は、2個の代表値 $a_1$ ,  $a_2$ を含む所定の範囲内にある相対値 $z_3(f, m)$ に対してハイレベル値(例えば1)をとり、この所定の範囲内でない代表値 $a_3$ に対してローレベル値(例えば0)をとるバイナリマスクの例である。この図の縦軸はバイナリマスクのゲインを示し、横軸は相対値 $z_3(f, m)$ (信号の到来方向(deg.))。この図に示すように、この例のバイナリマスクのハイレベル値はフラットであり、このハイレベル値とローレベル値とは不連続である。

[0098] なお、その他の構成や処理については、第1及び第2の実施の形態と同様である。すなわち、本形態では、第1及び第2の実施の形態で使用した滑らかな形状のマスク $M_{DC}(f, m)$ の代わりにバイナリマスク $B(f, m)$ を用い、周波数領域の信号値から、 $V$ 個の信号源から発せられた信号からなる混合信号(本形態ではこれを「限定信号」と呼ぶ)の値を抽出し、第1或いは第2の実施の形態の処理を実行する。

また、バイナリマスク $B(f, m)$ を用い、周波数領域の信号値から、 $V$ 個の信号源から発せられた信号からなる混合信号の値を抽出する処理は、周波数領域の観測信号値 $X_j(f, m)$ にバイナリマスク $B(f, m)$ を乗じることにより行う( $X^{\wedge}(f, m) = B(f, m) X(f, m)$ )。

[0099] <本形態の特徴>

従来の「従来法2:信号のスパース性を利用した方法」で分離性能を上げた際に分

離信号の歪が大きくなるのは、分離性能を上げるために上述の式(3)の  $\varepsilon$  を十分小さくすると、このバイナリマスクによって取り出される信号成分が制限され、本来原信号の成分として取り出されるべきサンプルの多くが取り出されないからである。つまり、この場合、各分離信号に多くの0成分が詰めこまれ、各分離信号を不連続とし、Musical Noiseを発生させることになる。

- [0100] これに対し、この形態では、バイナリマスク  $B(f, m)$  によって、2個以上  $M$  個以下の原信号からなる混合信号(限定信号)を抽出する。そのため、1個のみの信号の値を抽出する[従来法2]のバイナリマスクよりも、広い範囲の相対値  $z(f, m)$  に対する信号(サンプル)を限定信号として抽出できる。例えば、図12Aの例の場合、相対値  $z(f, m)$  が代表値  $a_1$  や  $a_2$  の近傍となるサンプル値のみではなく、相対値  $z(f, m)$  が  $a_1$  と  $a_2$  との間に位置するようなサンプル値も抽出できる。また、例えば、 $a_1$  と  $a_2$  との間に位置するようなサンプルは、代表値  $a_1$  或いは  $a_2$  に対応するサンプルである可能性が高い。
- [0101] そのため、ある時刻において同一周波数の観測信号が2個以上存在し、サンプル値が、本来対応すべき代表値から離れてしまった場合であっても、このようなサンプル値を抽出できる可能性は高くなる。その結果、分離信号に不連続に0成分が詰めこまれることによる品質劣化(Musical Noiseの発生)を抑制できる。

[0102] <バイナリマスクによる0詰めの影響の検証>

以下に、3人の話者による音声信号  $s_1, s_2, s_3$  が、2つの無指向性マイクで観測された場合 ( $N=3, M=2$  の場合) について、バイナリマスクによる0詰めの影響を議論する。

バイナリマスクによって0が詰められ失われた信号のパワーの比率を、

[数27]

$$\frac{\sum_t |s_i(t)|^2 - \sum_t |y_i(t)|^2}{\sum_t |s_i(t)|^2} \times 100 \dots (26)$$

と定義すると、従来の「信号のスパース性を利用した方法(従来法2)」では、 $s_1$ :17%、 $s_2$ :14%、 $s_3$ :23%もの信号のパワーがバイナリマスクにより失われた。

- [0103] 一方、本形態のバイナリマスク  $B(f, m)$  による信号のパワー劣化は、限定信号を  $s_1$  と  $s_2$  の2つの信号の混合とした場合には  $s_1$ :2.5%、 $s_2$ :5.7%であり、 $s_2$  と  $s_3$  の2つの信

号の混合とした場合には $s_2:8.1\%$ 、 $s_3:0.7\%$ であった。

すなわち、この形態では、バイナリマスク $B(f, m)$ による信号の劣化が従来法に比べて少ないことが分かる。これは、この形態では、Musical Noiseが発生しにくいことを示している。

[0104] <性能比較>

以下に、本形態のシミュレーション結果を示す。

[表5]

	SIR1	SIR2	SIR3	SDR1	SDR2	SDR3
従来法 2	15.4	10.3	14.6	9.8	11.9	9.2
本形態	-	8.4	16.4	-	15.0	20.9
	13.1	8.2	-	17.4	13.8	-

この例は、本形態のバイナリマスクで限定信号を抽出し、その限定信号にICAを提供して信号分離を行った例である。またこの例では、原信号として、3人の話者(男性2名・女性1名)による音声信号を用い、残響の無い環境でのこれらの混合信号を、2つの無指向性マイクで観測する状況をシミュレートしている。この表に示すように、この形態の方法では、分離性能SIRをほとんど落とすことなく、従来法2に比べ格段に高いSDRを得ることができる。これは、この形態の方法が、格段に低い歪で信号の分離を行っていることを示している。

[0105] [第7の実施の形態]

本形態も第1の本発明に係る実施の形態であり、上述の第6の実施の変形例である。すなわち、本形態も $2 \leq V \leq M$ の場合にバイナリマスクを用いて限定信号を抽出する形態であるが、バイナリマスク $B(f, m)$ の作成方法及び限定信号の算出処理に違いがある。以下では、このバイナリマスク $B(f, m)$ の作成方法、及び限定信号の算出処理にのみについて説明を行い、その他の処理及び機能構成については、第1の実施の形態或いは第2の実施の形態と同一であるため、説明を省略する。

[0106] この形態のバイナリマスク $B(f, m)$ は、上述の限定信号以外の観測信号成分を抽出するためのものである。すなわち、この形態のマスク作成部が作成するバイナリマスク $B(f, m)$ は、 $V$ 個の代表値(この集合を $G_k$ とする)を含む所定の範囲内にある相対値に対してローレベル値をとり、この所定の範囲内にない代表値( $G_k^c$ )に対してハイレベ

ル値をとり、ハイレベル値からローレベル値への推移が不連続な関数である。ただし、 $2 \leq V \leq M$ である。

- [0107] すなわち、この形態のマスク作成部51-kは、例えば、 $G_k^{\circ}$ に含まれる代表値について、上述の式(3)で示されるバイナリマスクを生成する。また、本形態でも、位相差 $z_1(f, m)$ 、振幅比 $z_2(f, m)$ 、位相差 $z_1(f, m)$ から得られる信号の到来方向 $z_3(f, m)$ 等を、相対値 $z(f, m)$ として使用することができる。図12Bは、本形態のバイナリマスク $B(f, m)$ の例示である。この例は、 $V=2$ 個の代表値 $a_1, a_2$ を含む所定の範囲内にある相対値 $z_3(f, m)$ に対してローレベル値(例えば0)をとり、この所定の範囲内にない代表値 $a_3$ に対してハイレベル値(例えば1)をとるバイナリマスクの例である。この図の縦軸はバイナリマスクのゲインを示し、横軸は相対値 $z_3(f, m)$ (信号の到来方向(deg.))。この図に示すように、この例のバイナリマスクのハイレベル値はフラットであり、このハイレベル値とローレベル値とは不連続である。

- [0108] また、この形態の限定信号抽出部は、周波数領域の信号値 $X_j(f, m)$ にこのバイナリマスク $B(f, m)$ を乗じた値を、周波数領域の信号値 $X_j(f, m)$ から減算し、限定信号値 $\hat{X}(f, m)$ を抽出する。例えば、上述の式(3)で示されるバイナリマスク $M_i(f, m)$ を集合 $G_k^{\circ}$ に含まれる $N-M$ 個の代表値について作成し、

[数28]

$$\hat{X}(f, m) = X(f, m) - \sum_{G_k^{\circ}} \{M_i(f, m)X(f, m)\} \dots (27)$$

を計算することで、 $M$ 個の原信号のみからなる限定信号の値 $\hat{X}(f, m)$ を算出する。なお、上述の式(3)のバイナリマスク $M_i(f, m)$ は、それぞれ1つの代表値のみに対してハイレベル値をとるバイナリマスクであるが、2つ以上の代表値に対してハイレベル値をとるバイナリマスクを用いて本形態の処理を実行してもよい。また、バイナリマスクの変わりに上述した滑らかな形状のマスクを用いて本形態の処理を実行してもよい。

限定信号 $\hat{X}(f, m)$ が算出されると、以後第1の実施の形態或いは第2の実施の形態と同様な限定信号分離、時間領域変換、信号統合の処理が行われる。

- [0109] [第8の実施の形態]

本形態は、第2の本発明に係る例であり、 $M$ 個のセンサで信号が観測される状況に

において、観測値をM次元領域でクラスタリングし、マスクを定義する。なお、以下では第1の実施の形態との相違点を中心に説明し、第1の実施の形態と共通する事項については説明を省略する。

図13は、本形態における代表値生成部430、マスク制御部40及び限定信号作成部450-kの構成を例示したブロック図である。なお、この図はV個の分離信号を得る1系統のみを示している。なお、本形態では $1 \leq V \leq M$ である。

- [0110] 本形態の信号分離装置と第1の実施の形態の信号分離装置1との構造上の相違点は代表値生成部及び限定信号作成部である。すなわち、第1の実施の形態の信号分離装置1の代表値生成部30(図1)の代わりに代表値生成部430(図13)が設けられ、信号分離装置1の限定信号作成部50-k(図1)の代わりに限定信号作成部450-k(図13)が設けられる。その他の構成については第1の実施の形態と同様である。

図14は、本形態における信号分離処理を説明するためのフローチャートである。以下、このフローチャートに添って、本形態の信号分離処理について説明する。

- [0111] まず、前処理として、各センサにおいて観測された時間領域の各観測信号 $x_j(t)$  ( $j=1, \dots, M$ )を記憶部2(図1)に格納しておく。そして、信号分離処理が開始されると、信号分離プロセッサ3は制御部10の制御のもと以下の処理を実行する。

まず信号分離プロセッサ3は、制御部10の制御のもと記憶部2にアクセスし、そこから各観測信号値 $x_j(t)$ を順次読み込み、周波数領域変換部20に送る(ステップS21)。周波数領域変換部20は、短時間離散フーリエ変換等によって、これらの信号値を時間ごとの周波数領域の観測信号値 $X_j(f, m)$ に順次変換し、一時記憶部90に格納する(ステップS22)。

- [0112] 次に、クラスタリング部432(図13)が、一時記憶部90(図1)に格納された周波数領域の観測信号値 $X_1(f, m), \dots, X_M(f, m)$ を読み出す。そして、クラスタリング部432(図13)は、これら周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ からなる観測信号ベクトル(「第1のベクトル」に相当) $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$ を、周波数fごとにN個ずつのクラスタ $C_i(f)$  ( $i=1, \dots, N$ )にクラスタリングし、信号源数Nと等しいN個のクラスタ $C_i$  ( $i=1, 2, \dots, N$ )を生成する(ステップS23)。なお生成



されたN個のクラスタ $C_i$ は、一時記憶部90(図1)に格納される。

- [0113] ここで本形態におけるクラスタとは、観測信号ベクトル $X(f, m)$ の集合であり、離散時間 $m$ の集合 $T_i$ を用いて $C_i(f) = \{X(f, m) \mid m \in T_i\}$ と表記される。また、クラスタリングの目的は、同じ信号源が支配的である(主な成分を持つ)サンプル(観測信号ベクトル $X(f, m)$ )を同じクラスタに分類することである。なお、得られるN個のクラスタ $C_1(f), \dots, C_N(f)$ は、必ずしも、互いに素( $C_i(f) \cap C_j(f)$ が空集合,  $i \neq j$ )である必要はなく、またクラスタに属さない要素

[数29]

$$X(f, m) \notin \bigcup_{i=1}^N C_i$$

が存在してもよい。

- [0114] [クラスタリング部432での処理の詳細]

ここでクラスタリング部432の処理をさらに詳細に説明する。

この例のクラスタリング部432は、クラスタリングを適切に実行できるように、すなわち同じ信号源が支配的であるサンプル(観測信号ベクトル $X(f, m)$ )が同じクラスタに分類されるように、各サンプルの正規化を行ってからクラスタリングを行う。

具体的には、例えばまず正規化部432a(図13)が、一時記憶部90(図1)から観測信号ベクトル $X(f, m)$ を読み込み、

[数30]

$$\text{sign}(X_j(f, m)) = \begin{cases} X_j(f, m) / |X_j(f, m)| & (|X_j(f, m)| \neq 0) \\ 0 & (|X_j(f, m)| = 0) \end{cases} \quad \dots(28)$$

を算出し、

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \text{sign}(X_j(f, m)) & (|X_j(f, m)| \neq 0) \\ X(f, m) & (|X_j(f, m)| = 0) \end{cases} \quad \dots(29)$$

の正規化を行い、クラスタ生成部432bが、この正規化結果のクラスタリングを行う。

- [0115] また、さらに必要であれば、この例の正規化部432aは、式(28)(29)の正規化を行った後、さらに、

[数31]

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \|X(f, m)\| & (\|X(f, m)\| \neq 0) \\ X(f, m) & (\|X(f, m)\| = 0) \end{cases} \quad \dots(30)$$

の正規化を行い、クラスタ生成部432bが、この正規化結果のクラスタリングを行う。ただし、ベクトルの長さ $\|X(f, m)\|$ は $X(f, m)$ のノルムであ

り、具体的には、例えば、

[数32]

$$L_k(X(f, m)) = \left( \sum_{j=1}^M |X_j|^k(f, m) \right)^{1/k} \quad \dots(31)$$

で定義される $L_2$ ノルム $\cdot X(f, m) \cdot$  ( $\|X(f, m)\|$ )を用いる。

[0116] また、クラスタ生成部432bが行うクラスタリングの方法としては、例えば、階層的クラスタリングやk-meansクラスタリング等の多くの教科書で説明されている方法を用いることができる(例えば、「尾上守夫 監訳“パターン識別”，新技術コミュニケーションズ，ISBN 4-915851-24-9，第10章」等参照。)。なお、いずれのクラスタリング方法も、2つのサンプル $X(f, m)$ と $X'(f, m)$ の距離が定義され、それに従ってサンプル間の近さが測られ、なるべく距離の近いサンプル同士が同じクラスタに含まれるようにクラスタリングするものである。

[0117] 例えば、上述の式(29)のみによってサンプルが正規化された場合、クラスタ生成部432bは、正規化された2つの観測信号ベクトル $X(f, m)$ 間のコサイン距離を距離尺度として用いてクラスタリングを行う。なお、2つのサンプル $X(f, m)$ と $X'(f, m)$ のコサイン距離は、

$$1 - X^H(f, m) \cdot X'(f, m) / (\|X(f, m)\| \cdot \|X'(f, m)\|) \quad \dots(32)$$

によって定義される。

[0118] また、上述の式(29)と式(30)によってサンプルが正規化された場合、クラスタ生成部432bは、正規化された2つの観測信号ベクトル間の差 $(X(f, m) - X'(f, m))$ の $L_2$ ノルム $\cdot X(f, m) - X'(f, m) \cdot$  ( $\|X(f, m) - X'(f, m)\|$ )や、任意の $k$ による $L_k$ ノルム、或いはコサイン距離(式(32))を距離尺度として用いてクラスタリングを行う([クラスタリング部432での処理の詳細]の説明終わり)。

次に、代表値計算部433が、一時記憶部90(図1)に格納された各クラス $C_i(f)$ を順

次抽出し、各クラス $C_i(f)$ を代表する代表ベクトル(「第2のベクトル」に相当) $a_i(f)$ を算出する(ステップS24)。

[0119] [代表値計算部433での処理の詳細]

例えば、まず代表値計算部433の代表ベクトル生成部433a(図13)が、一時記憶部90(図1)に格納された各クラス $C_i(f)$ を順次抽出し、各クラス $C_i(f)$ に属するサンプルの値 $X(f, m)$ の平均値

[数33]

$$a_i(f) = \sum_{X(f,m) \in C_i(f)} X(f,m) / |C_i(f)|$$

を各信号源に関する代表ベクトル $a_i(f)$ として算出する。或いは、各クラス $C_i(f)$ に属するサンプル $X(f, m)$ を適度に量子化し、最頻値を求めてこれを代表ベクトル $a_i(f)$ としてもよい。このように求められた代表ベクトル $a_i(f)$ は一時記憶部90(図1)に格納される。

[0120] 次に、並び替え部433b(図13)が、一時記憶部90(図1)から、これらの代表ベクトル $a_1(f), \dots, a_N(f)$ を読み出し、これらの各代表ベクトル $a_1(f), \dots, a_N(f)$ の各源信号 $s_k(t)$ との対応が、すべての周波数 $f$ において等しくなるように、各代表ベクトル $a_i(f)$ の添字 $i$ を付け替える(ステップS25)。

具体的には、例えばまず、並び替え部433b(図13)が、読み出した各周波数 $f$ の代表ベクトル $a_i(f)$ を用い、

[数34]

$$\theta_i(f) = \arccos \frac{\arg(a_{ji}(f)/a_{ji}(f))}{2\pi f v^{-1} \|d_j - d_j\|} \dots (33)$$

の演算によって、各周波数 $f$ に対する源信号 $i$ の到来方向の推定値 $\theta_i(f)$ を算出する。なお、 $d_j$ はセンサ $j$ の位置、 $v$ は信号の速さ、 $a_{ji}(f)$ は代表ベクトル $a_i(f)$ の $i$ 番目の要素であり、 $d_j$ 及び $v$ は、例えば、予め一時記憶部90に格納されているデータを用いることとする。

[0121] 算出された各推定値 $\theta_i(f)$ は、例えば、その算出に用いた代表ベクトル $a_i(f)$ に対応付けられて一時記憶部90(図1)に格納される。次に、並び替え部433b(図13)は

、例えば、一時記憶部90から各推定値  $\theta_i(f)$  を読み込み、これらを各周波数  $f$  ごとに所定の順序(例えば、昇順、降順等)で並び替える。なお、この並び替えは、例えば公知の並び替えアルゴリズムによって行われる。そして、この並び替え後の各  $f$  における各代表ベクトル  $a_i(f)$  の順番を示す情報  $(j'(f, a_i(f)) = 1, 2, \dots, N)$  が、一時記憶部90(図1)に格納される。そして、並び替え部433b(図13)は、例えば、この順序情報  $j'(f, a_i(f))$  を一時記憶部90から読み込み、当該  $a_i(f)$  が  $j'(f, a_i(f))$  番目の源信号に対応するとして、各代表ベクトルと  $i$  との対応付けを変更する( $a_i(f)$  の添字  $i$  を付け替える)。そして、この添字  $i$  が付け替えられた各代表ベクトル  $a_i(f)$  は、一時記憶部90(図1)に格納される。

- [0122] 次に、マスク制御部40は、これら各代表ベクトル  $a_i(f)$  を要素に持つ集合  $G_0$  を特定するデータを変数  $SG_0$  に代入し、この変数  $SG_0$  を一時記憶部90に格納する。また、マスク制御部40は、集合  $G$  を特定する変数  $SG$  を  $G = \phi$  (空集合)に初期化し、変数  $k$  を0とし、それらを一時記憶部90に格納する(ステップS26)。

次に、マスク制御部40での制御のもと、 $N$ 個すべての分離信号が得られるまで、限定信号作成部50- $k$  ( $k=1, \dots, u$ )、限定信号分離部60- $k$  及び時間領域変換部70- $k$  の複数系統( $u$ 系統)による処理が行われる。

- [0123] まず、マスク制御部40は、一時記憶部90に格納された変数  $k$  に1を加えた値を新たな変数  $k$  とし、再び一時記憶部90に格納する(ステップS27)。

次に、マスク制御部40は、一時記憶部90(図1)から変数  $SG_0$  及び  $SG$  を呼び出す。そして、マスク制御部40は、変数  $SG_0$  によって特定される集合  $G_0$  から、 $SG$  によって特定される集合  $G$  の補集合 ( $G^c$  ( $\alpha^c$  は  $\alpha$  の補集合を示す)) の元を含む適当な  $V$  ( $\leq M$ ) 個の代表ベクトル  $a_p(f)$  ( $p=1, \dots, V$ ) (「第3のベクトル」に相当)の集合  $G_k$  を選択し、この集合  $G_k$  を特定するデータを変数  $SG_k$  に代入し、この変数  $SG_k$  を一時記憶部90に格納する(ステップS28)。すなわち、マスク制御部40は、各代表ベクトル  $a_1(f), \dots, a_N(f)$  の中から、限定信号として取り出す  $V$  個の信号に対応する  $V$  個の代表ベクトル  $a_p(f)$  ( $p=1, \dots, V$ ) を抽出する。

- [0124] 本形態では、この集合  $G_k$  に含まれる代表ベクトル  $a_p(f)$  に近いサンプル値  $X(f, m)$  を抽出し、集合  $G_k$  に含まれない代表ベクトル(集合  $G_k^c$  の要素、 $*^c$  は  $*$  の補集合を

示す)に近いサンプル値 $X(f, m)$ を抽出しないことで、 $V$ 個の信号が混合した限定信号 $\hat{X}(f, m)$ を作成する。

そのために、本形態では、限定信号作成部450-kのマスク作成部451-k(図13)が、一時記憶部90(図1)から変数 $SG_k$ 、 $SG_0$ 及び観測信号ベクトル $X(f, m)$ を読み込み、以下のマスク $M_k(f, m)$ を生成する(ステップS29)。

[0125] [数35]

$$M_k(f, m) = \begin{cases} 1 & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

なお、 $D(X(f, m), a_i(f))$ は、ベクトル $X(f, m)$ と $a_i(f)$ とのマハラノビス平方距離

$$D(X(f, m), a_i(f)) = (X(f, m) - a_i(f))^H \Sigma^{-1} (X(f, m) - a_i(f)) \text{ を示し、}$$

$$\Sigma \text{ はクラス } C_i \text{ の共分散行列 } \Sigma = \frac{1}{|C_i|} \sum (X(f, m) - a_i(f))(X(f, m) - a_i(f))^H$$

を示し、 $|C_i|$ は、クラス $C_i$ に属するサンプル数を示す。また、源信号の大きさがほぼ同じであることが分かっている場合、共分散行列 $\Sigma = I$ (単位行列)としてもよい。

[0126] このマスク $M_k(f, m)$ は、一時記憶部90(図1)に格納され、限定信号抽出部452-k(図13)は、一時記憶部90からマスク $M_k(f, m)$ と観測信号ベクトル $X(f, m)$ とを読み込み、マスク $M_k(f, m)$ と観測信号ベクトル $X(f, m)$ との積

$$\hat{X}_k(f, m) = M_k(f, m) \cdot X(f, m)$$

を演算し、 $V$ 個の信号源から発せられた限定信号値 $\hat{X}_k(f, m)$ を抽出する(ステップS30)。

[0127] この限定信号値 $\hat{X}_k(f, m)$ は、一時記憶部90(図1)に格納された後、限定信号分離部60-kに送られ、限定信号分離部60-kは、この限定信号値 $\hat{X}_k(f, m)$ を用い、限定信号の信号分離を行う(ステップS31)。ここで、限定信号値 $\hat{X}_k(f, m)$ は、 $V(1 \leq V \leq M)$ 個の信号源から発せられた信号によって構成された混合信号の値であると近似される。よって、その分離行列の推定には[従来法1]で述べた独立成分分析による方法を利用できる。すなわち独立成分分析の入力として、観測信号値 $X$ の代わりに限定信号値 $\hat{X}_k(f, m)$ を用い、例えば[従来法1]で述べた式(2)を用いて分離

を行う。なお、 $V=1$ の場合は、ステップS31の処理は不要である。

- [0128] 本実施例におけるICAによる分離では、まず、ICA分離行列推定部61-k(図2)において、限定信号値 $\hat{X}_k(f, m)$ を用い、前述の式(2)の学習則に従い分離行列 $W(f, m)$ を生成し、この分離行列 $W(f, m)$ を一時記憶部90に格納する。なお、この分離行列 $W(f, m)$ の生成には、例えば、以下のパーミュテーション・スケーリング解決部62-kからの出力値 $Y_k(f, m)$ のフィードバックを用いる。生成された分離行列 $W(f, m)$ はパーミュテーション・スケーリング解決部62-kに送られる。パーミュテーション・スケーリング解決部62-kは、この分離行列 $W(f, m)$ と限定信号値 $\hat{X}_k(f, m)$ を用い、 $Y_k(f, m) = W(f, m) \hat{X}_k(f, m)$ の演算を行い、それぞれの分離信号値 $Y_k(f, m) = [Y_{k1}^{\Pi_{k1}}(f, m), \dots, Y_{kV}^{\Pi_{kV}}(f, m)]^T$ を生成し、それを一時記憶部90に格納する。そして、パーミュテーション・スケーリング解決部62-kは、例えば、この分離信号値 $Y_k(f, m)$ をフィードバックし、[従来法1]で述べた方法でPermutation問題を解決する。Permutation問題の解決後、パーミュテーション・スケーリング解決部62-kは、さらに、分離信号値 $Y_{kq}$  ( $q=1, \dots, V$ )がどの源信号に対応するのかを示すタグ $\Pi_{kq}$ を、分離信号値 $Y_{kq}$  ( $q=1, \dots, V$ )に付与し、これらに対応付けて一時記憶部90に格納する。ここでは、このタグ $\Pi_{kq}$ を分離信号値 $Y_{kq}$ の上付添字 $\Pi_{kq}$ として表記する。

- [0129] 具体的には、例えば、パーミュテーション・スケーリング解決部62-kが、一時記憶部90から抽出した分離行列 $W(f)$ の逆行列( $N \neq M$ の場合はMoore-Penrose型擬似逆行列)を用い、

[数36]

$$\theta_q(f) = \arccos \frac{\arg([W_{jq}^{-1}(f)]/[W_{jq}^{-1}(f)])}{2\pi f v^{-1} \|d_j - d_j\|} \dots (34)$$

(ここで $v$ は信号の速さ、 $d_j$ はセンサ $j$ の位置)

の演算によって得られる信号の推定到来方向 $\theta_q(f)$ と、一時記憶部90から抽出した変数 $SG_k$ が示す集合 $G_k$ に含まれる代表ベクトル $a_p(f)$ とを比較し、 $\theta_q$ に最も近い代表ベクトル $a_p(f)$ を $q$ 番目の分離信号 $Y_{kq}$ に対応付ける(ステップS32)。つまり、パーミュテーション・スケーリング解決部62-kは、この分離信号 $Y_{kq}$ に対し、代表値 $a_i$ を示すタ

タグ  $\Pi_{kq}$  を付与する(対応付ける)。

- [0130] この後、パーミュテーション・スケーリング解決部62-kが、一時記憶部90から分離行列  $W(f)$  を抽出し、その各行  $w_q(f)$  を

$$w_q(f) \leftarrow [W^{-1}(f)]_{jq} w_q(f)$$

とすることにより、ICAのスケーリング問題を解決し、スケーリング問題解決後の分離行列  $W(f)$  を一時記憶部90に格納する。なお、後の信号統合部80における処理のため、この処理ではすべての系列  $k$  において同じ  $j$  を用いることが望ましい。

タグ  $\Pi_{kq}$  が付与された各分離信号値  $y_{kq}$  は、時間領域変換部70-kに送られる。時間領域変換部70-kは、例えば、短時間逆離散フーリエ変換等により、時間周波数領域で得られている各分離信号値  $y_{kq}$  を時間領域の信号値に変換し、その変換値を一時記憶部90に格納する。(ステップS33)。なお、これら時間領域の信号値  $y_k(t) = [y_{k1}^{\Pi_{k1}}(t), \dots, y_{kV}^{\Pi_{kV}}(t)]^T$  にも上述のタグ  $\Pi_{kq}$  が関連付けられる。この関連付けを行う場合、まず、時間領域変換部70-kが、一時記憶部90から、周波数領域の信号値  $Y_{kq}$  に対応付けられているタグ  $\Pi_{kq}$  を各周波数及び時間について抽出する。次に、時間領域変換部70-kは、各周波数及び時間におけるタグ  $\Pi_{kq}$  がすべて等しいか否かを判断する。ここでこれらがすべて等しかった場合には、時間領域の信号値  $y_k$  のタグとして、周波数領域の信号値  $Y_{kq}$  に対応付けられているタグ  $\Pi_{kq}$  を対応付ける。一方、これらがすべて等しくなかった場合には、多数決にて時間領域の信号値  $y_{kq}$  のタグを決定する。

- [0131] 次に、マスク制御部40において、一時記憶部90から変数  $SG$  と  $SG_k$  とを抽出し、これらが示す  $G$  と  $G_k$  との和集合  $G \cup G_k$  を新たな集合  $G$  とし、この集合  $G$  を変数  $SG$  に代入し、この変数  $SG$  を一時記憶部90に格納する(ステップS34)。また、マスク制御部40は、一時記憶部90から変数  $SG$  と  $SG_0$  とを読み出し、この新たな集合  $G$  が集合  $G_0$  と等しいか否かを判断する(ステップS35)。ここで、 $G = G_0$  でなければステップS27の処理に戻る。

- [0132] 一方、 $G = G_0$  であれば、信号統合部80において、一時記憶部90から各系統  $k$  (時間領域変換部70-k /  $k = 1, \dots, u$ ) から出力された分離信号  $y_{kp}(t)$  を読み出し、これらの選択／統合を行い、 $N$  個すべての分離信号を得る(ステップS36)。具体的に

は、例えば、まず信号統合部80は、一時記憶部90から読み出した各分離信号 $y_{kp}(t)$ のタグ $\Pi_{kq}$ を比較する。ここで、複数の系統 $k$ において同じタグを持つ分離信号値 $y_{kq}(t)$ が存在しないと判断された場合、信号統合部80は、すべての分離信号値 $y_{kq}(t)$ を最終的な分離信号値 $y_i(t)$  ( $i=1, \dots, N$ )として出力する(ステップS37)。一方、複数の系統において同じタグを持つ分離信号値が存在すると判断された場合、信号統合部80は、これらのタグが等しい分離信号値のどれか1つを適当に選択し、最終的な分離信号値 $y_i(t)$ として出力するか、同じタグを持つ分離信号値の平均を計算し、それを出力信号とする(ステップS37)。

[0133] ここで、どれか1つの分離信号値 $y_{kq}(t)$ を適当に選択し、最終的な分離信号値 $y_i(t)$ として出力する処理の場合、信号統合部80は、例えば、同じタグ $a_i$ を持つ分離信号値 $y_{kq}(t)$ の中で最大パワーを持つものを最終的な分離信号値 $y_i(t)$ として出力する。また、同じタグを持つ分離信号値の平均を最終的な分離信号値 $y_i(t)$ として出力する処理の場合、信号統合部80は、例えば、

[数37]

$$y_i(t) = \frac{1}{K} \sum_{\Pi_{kq}=a_i} y_{kq}(t)$$

( $K$ は同じタグ $a_i$ を持つ分離信号の個数)

とする。以上により、 $N$ 個の信号が少ない歪で分離される。

[0134] なお、本形態の変形として、マスク $M(f, m)$ を生成せず、

[数38]

$$\hat{X}_k(f, m) = \begin{cases} X(f, m) & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

として、直接限定信号値を生成してもよい。すなわち、例えば、限定信号作成部450-kが、観測信号ベクトル $X(f, m)$ に対し、

[数39]

$$\max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f))$$

を満たすか否かを判断し、満たすと判断した観測信号ベクトル $X(f, m)$ を、信号源か



ら発せられた信号の値として抽出することとしてもよい。

[0135] 〔第9の実施の形態〕

本形態は、第3の本発明に係る実施の形態である。

<構成>

図15は、本形態におけるブランド信号分離装置500の構成を例示したブロック図である。なお、この図における矢印はデータの流れを示すが、制御部521や一時記憶部522に出入りするデータの流れは省略してある。すなわち、データが制御部521や一時記憶部522を経由する場合であっても、その経由の過程は省略してある。

[0136] まず、この図を用いて、本形態の構成について説明する。

図15に例示するように、本形態の信号分離装置500は、記憶部501とこれに有線或いは無線で電氣的に接続された信号分離プロセッサ502とを有している。

記憶部501は、例えば、ハードディスク装置、フレキシブルディスク、磁気テープ等の磁気記録装置、DVD-RAM (Random Access Memory)、CD-R (Recordable) / RW (ReWritable) 等の光ディスク装置、MO (Magneto-Optical disc) 等の光磁気記録装置、EEP-ROM (Electrically Erasable and Programmable-Read Only Memory)、フラッシュメモリ (flash memory) 等の半導体メモリ等である。また、記憶部501は、信号分離プロセッサ502と同一の筐体内に存在してもよいし、別個の筐体に構成されてもよい。

[0137] またこの例の信号分離プロセッサ502は、例えばプロセッサやRAM等によって構成されるハードウェアであり、周波数領域変換部511、混合行列推定部512、パーミューテーション問題解決部513、スケーリング問題解決部514、列選択部516、行列生成部517、分離行列生成部518、分離信号生成部519、時間領域変換部520、制御部521及び一時記憶部522を有している。また、この例の混合行列推定部512は、クラスタリング部512a、代表ベクトル計算部512b及びベクトル統合部512cを有している。さらに、クラスタリング部512aは、正規化部512aa及びクラスタ生成部512abを有している。

[0138] <処理>

図16は、本形態における信号分離装置500の処理の全体を説明するためのフロ

一チャートである。以下、図15及び図16を用いて、信号分離装置500の処理を説明していく。なお、以下では $N$  ( $N \geq 2$ ) 個の信号源から発せられた信号が混合され、 $M$  個のセンサで観測された場合について説明する。

[0139] [処理の全体]

信号分離装置500は、制御部521の制御のもと、以下の処理を実行する。

まず、 $M$  個のセンサで観測された観測信号の値  $x_1(t), \dots, x_M(t)$  ( $t$  は時間) が、記憶部501から読み込まれ、周波数領域変換部511に入力される(図15)。周波数領域変換部511は、これらの観測信号値  $x_1(t), \dots, x_M(t)$  を、短時間離散フーリエ変換等により、周波数領域の信号値(周波数  $f$  ごとの時系列データ)  $X_1(f, m), \dots, X_M(f, m)$  ( $m$  は離散時間) に変換する(ステップS51)。これら周波数領域の信号値  $X_1(f, m), \dots, X_M(f, m)$  は、一時記憶部522に格納され、混合行列推定部512のクラスタリング部512aによって読み込まれる。クラスタリング部512aは、これらによって構成される観測信号ベクトル  $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]^T$  を、周波数  $f$  ごとに  $N$  個ずつのクラス  $C_i(f)$  ( $i = 1, \dots, N$ ) にクラスタリングする(ステップS52)。各クラス  $C_i(f)$  は、代表ベクトル計算部512bに送られ、代表ベクトル計算部512bは、各クラス  $C_i(f)$  の代表ベクトル  $a_i(f)$  を算出する(ステップS53)。各代表ベクトル  $a_i(f)$  は、一時記憶部522に格納され、ベクトル統合部512cは、これらを順次抽出し、各代表ベクトル  $a_i(f)$  を列とする推定混合行列  $A(f) = [a_1(f), \dots, a_N(f)]$  を生成する(ステップS54)。生成された推定混合行列  $A(f)$  は、一時記憶部522に格納される。

[0140] パーミュテーション問題解決部513は、推定混合行列  $A(f)$  を一時記憶部522から読み込み、推定混合行列  $A(f)$  の列を並び替えてパーミュテーション問題を解決する(ステップS55)。なお、この処理には、後述する分離信号の値  $Y_1(f, m), \dots, Y_N(f, m)$  をフィードバックして用いることも可能であり、その場合、より正確にパーミュテーション問題を解決できる。

[0141] 次に、スケーリング問題解決部514において推定混合行列  $A(f)$  の列を正規化してスケーリング問題を解決した後(ステップS56)、この推定混合行列  $A(f)$  を用いて、分離行列生成部518が分離行列  $W(f, m)$  を生成する(ステップS57)。生成された分離行列  $W(f, m)$  は、一時記憶部522に格納された後、そこから分離信号生成部519

に送られ、分離信号生成部519は、一時記憶部522から周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ を読み込み、 $Y(f, m) = W(f, m) X(f, m)$ の演算により、分離信号ベクトル $Y(f, m) = [Y_1(f, m), \dots, Y_N(f, m)]^T$ を算出する(ステップS58)。算出された分離信号値 $Y_1(f, m), \dots, Y_N(f, m)$ は、一時記憶部522に格納され、パームミュレーション問題解決部513にフィードバックされる他、時間領域変換部520にも送られる。そして、時間領域変換部520は、分離信号値 $Y_1(f, m), \dots, Y_N(f, m)$ を、添字 $i$ ごとの短時間逆フーリエ変換等により時間領域の信号値 $y_1(t), \dots, y_N(t)$ に変換し(ステップS59)、時間領域での分離信号値 $y_i(t)$ が得られる。

[0142] [混合行列推定部512での処理の詳細]

次に、混合行列推定部512での処理の詳細について説明する。なお、以下の処理は、周波数ごとに適用されるものである。

まず、クラスタリング部512aは、一時記憶部522から読み込んだすべてのセンサの観測信号成分 $X_1(f, m), \dots, X_M(f, m)$ をまとめ、これらを観測信号ベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]^T$ として関連付ける。そして、クラスタリング部512aは、クラスタリングによって信号源と等しい数 $N$ 個のクラスタ $C_i(f)$ を生成し、これらを一時記憶部522に格納する(ステップS52)。

[0143] ここでクラスタとは、観測信号ベクトル $X(f, m)$ の集合であり、離散時間 $m$ の集合 $T_i$ を用いて $C_i(f) = \{X(f, m) \mid m \in T_i\}$ と表記する。また、クラスタリングの目的は、同じ信号源が支配的である(主な成分を持つ)サンプル(観測信号ベクトル $X(f, m)$ )を同じクラスタに分類することである。なお、得られる $N$ 個のクラスタ $C_1(f), \dots, C_N(f)$ は、必ずしも、互いに素( $C_i(f) \cap C_j(f)$ が空集合,  $i \neq j$ )である必要はなく、またクラスタに属さない要素

[数40]

$$X(f, m) \notin \bigcup_{i=1}^N C_i$$

が存在してもよい。

[0144] 次に、代表ベクトル計算部512bは、一時記憶部522から各クラスタ $C_i(f)$ を読み込み、各クラスタ $C_i(f)$ に属するサンプル $X(f, m)$ の平均値

[数41]

$$a_i(m) = \sum_{X(f,m) \in C_i(f)} X(f,m) / |C_i(f)|$$

を各信号源に関する代表ベクトル $a_i(f)$ として算出する(ステップS53)。或いは、各クラス $C_i(f)$ に属するサンプル $X(f, m)$ を適度に量子化し、最頻値を求めてこれを代表ベクトル $a_i(f)$ としてもよい。

- [0145] 最後に、ベクトル統合部12cでN個の代表ベクトル $a_i(f)$ をまとめて、混合行列 $H(f) = [h_1(f), \dots, h_N(f)]$ の推定行列である推定混合行列 $A(f) = [a_1(f), \dots, a_N(f)]$ を生成して出力する(ステップS54)。なお、推定混合行列 $A(f)$ は、各ベクトルの順序に関する任意性(パーミュテーションの任意性)と、各ベクトルの大きさの任意性(スケールリングの任意性)を含んでいる。すなわち、代表ベクトル $a_i(f)$ は、 $h_{\Pi(i)}(f)$ に任意の複素数を掛けたものとして推定される。ここで、 $\Pi$ は、パーミュテーションの任意性を表現する順列である。

- [0146] [クラスタリング部512aでの処理の詳細]

次にクラスタリング部512aの処理をさらに詳細に説明する。

この例のクラスタリング部512aは、クラスタリングを適切に実行できるように、すなわち同じ信号源が支配的であるサンプル(観測信号ベクトル $X(f, m)$ )が同じクラスに分類されるように、正規化部512aaで各サンプルの正規化を行ってからクラスタリングを行う。

- [0147] 具体的には、この例の正規化部512aaは、

[数42]

$$\text{sign}(X_j(f,m)) = \begin{cases} X_j(f,m)/|X_j(f,m)| & (|X_j(f,m)| \neq 0) \\ 0 & (|X_j(f,m)| = 0) \end{cases} \quad \dots(35)$$

を算出し、

$$X(f,m) \leftarrow \begin{cases} X(f,m)/\text{sign}(X_j(f,m)) & (|X_j(f,m)| \neq 0) \\ X(f,m) & (|X_j(f,m)| = 0) \end{cases} \quad \dots(36)$$

の正規化を行った後にクラスタリングを行う。

また、さらに必要であれば、この例の正規化部512aaは、さらに、

[数43]

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \|X(f, m)\| & (\|X(f, m)\| \neq 0) \\ X(f, m) & (\|X(f, m)\| = 0) \end{cases} \quad \dots (37)$$

の正規化を行った後にクラスタリングを行う。ただし、ベクトルの長さ  $\|X(f, m)\|$  は  $X(f, m)$  のノルムであり、具体的には、例えば、

[数44]

$$L_k(X(f, m)) = \left( \sum_{j=1}^M |X_j|^k(f, m) \right)^{1/k} \quad \dots (38)$$

で定義される  $L_2$  ノルム  $\|X(f, m)\|_2 = \sqrt{X(f, m) \cdot X(f, m)}$  を用いる。

[0148] また、クラスタリングの方法としては、例えば、階層的クラスタリングやk-meansクラスタリング等の多くの教科書で説明されている方法を用いる(例えば、「尾上守夫 監訳 “パターン識別”, 新技術コミュニケーションズ, ISBN 4-915851-24-9, 第10章」等参照。)。なお、いずれのクラスタリング方法も、2つのサンプル  $X(f, m)$  と  $X'(f, m)$  の距離が定義され、それによってサンプル間の近さが測られ、なるべく距離の近いサンプル同士が同じクラスタに含まれるようにクラスタリングするものである。

[0149] 例えば、上述の式(36)のみによってサンプルが正規化された場合、クラスタリング部512aは、正規化された2つの観測信号ベクトル  $X(f, m)$  間のコサイン距離を距離尺度として用いてクラスタリングを行う。なお、2つのサンプル  $X(f, m)$  と  $X'(f, m)$  のコサイン距離は、

$$1 - X^H(f, m) \cdot X'(f, m) / (\|X(f, m)\| \cdot \|X'(f, m)\|) \quad \dots (39)$$

によって定義される。

[0150] また、上述の式(36)と式(37)によってサンプルが正規化された場合、クラスタリング部512aは、クラスタ生成部512abにおいて、上記の正規化された2つの観測信号ベクトル間の差  $(X(f, m) - X'(f, m))$  の  $L_2$  ノルム  $\|X(f, m) - X'(f, m)\|_2 = \sqrt{X(f, m) - X'(f, m) \cdot X(f, m) - X'(f, m)}$  や、任意のkによる  $L_k$  ノルム、或いはコサイン距離(式(39))を距離尺度として用いてクラスタリングを行う。

以上の操作により、各クラスタ  $C_i$  の代表ベクトル  $a_i(f)$  が混合ベクトル  $h_k(f)$  の推定(大きさの任意性を含む)となる理由を説明する。

[0151] クラスタ  $C_i$  には、ある源信号  $S_k$  のみが支配的で他の源信号は0に近いような観測信

号ベクトル $X(f, m)$ が集められている。この状況は、

$$X(f, m) = h_k(f) S_k(f, m) \cdots (40)$$

と近似できる。

そしてこの関係と式(36)の正規化により、

$$X \leftarrow X / \text{sign}(X) = h_k S_k / \text{sign}(H_{jk} S_k) = \text{sign}^*(H_{jk}) |S_k| h_k \cdots (41)$$

となる。なおここでは、 $\text{sign}(H_{jk} S_k) = \text{sign}(H_{jk}) \text{sign}(S_k)$ ,  $1 / \text{sign}(H_{jk}) = \text{sign}^*(H_{jk})$  ( $\cdot^*$ は複素数の共役をとる操作), 及び  $S_k / \text{sign}(S_k) = |S_k|$  なる関係を用いた。また、これらの記載において $f, m$ は省略してある。

[0152] また、式(37)の正規化と式(40)により、

$$X \leftarrow X / \cdot X \cdot = \text{sign}^*(H_{jk}) |S_k| h_k / (|S_k| \cdot \cdot h_k) = \text{sign}^*(H_{jk}) h_k / \cdot h_k \cdots (42)$$

となる。なおここでは、 $\cdot \text{sign}^*(H_{jk}) |S_k| h_k \cdot = |S_k| \cdot \cdot h_k$  なる関係を用いた。また、これらの記載においても $f, m$ は省略してある。

ここで、式(41)により、式(36)によって正規化された観測信号ベクトル $X(f, m)$ は、混合ベクトル $h_k(f)$ を $\text{sign}^*(H_{jk}(f))$ 倍したベクトルの直線上に集まることがわかる。そして、直線上のどこに乗るかは、信号源の大きさ $|S_k(f, m)|$ による。また、式(42)により、式(37)によって正規化された観測信号ベクトル $X(f, m)$ は、複素空間での1点 $\text{sign}^*(H_{jk}(f)) h_k(f) / \cdot h_k(f) \cdot$ に集まることがわかる。これらは、正規化された観測信号ベクトル $X(f, m)$ の平均として算出した代表ベクトル $a_i(f)$ が、大きさの任意性を含む混合ベクトル $h_k(f)$ の推定となっていることを示している。

[0153] [パーミュテーション問題解決部513での処理の詳細]

次に、パーミュテーション問題解決部513での処理の詳細について説明する。

パーミュテーション問題解決部513では、各周波数 $f$ で算出された推定混合行列 $A(f)$ の列の並べ替えを行い、同じ信号源 $s_k(t)$ に関する代表ベクトル $a_i(f)$ がすべての周波数 $f$ で同じになるようにする(ステップS55)。すなわち、各分離信号 $Y_1(f, m)$ ,  $\cdots$ ,  $Y_N(f, m)$ と各信号源との対応が各周波数 $f$ において同一となるように添字 $i$ を付け替える。そのために、例えば、従来の技術と同様、非特許文献2の手順に基づいて2種類の情報を用いる。

[0154] 1つ目の情報は、信号源の到来方向などの位置情報である。従来のICAを用いた方法では、分離行列WをICAにより求めて、そのムーア・ペンローズ(Moore-Penrose)型擬似逆行列 $W^+$ ( $M=N$ の場合は逆行列 $W^{-1}$ に一致)から位置情報を得ていた。ここで、このムーア・ペンローズ型擬似逆行列 $W^+$ は、混合行列の推定 $A(f)$ とみなせる。そのため、本形態では従来のICAを用いた方法とは異なり、推定混合行列 $A(f)$ そのものをムーア・ペンローズ型擬似逆行列 $W^+$ とみなし、その各列から直接位置情報を得ることができる。具体的には、例えば、

[数45]

$$\theta_i = \cos^{-1} \frac{\text{angle}(A_{ji}(f)/A_{ji}(f))}{2\pi f c^{-1} \|d_j - d_{j'}\|} \dots (43)$$

によって位置情報を得ることができる。ここで、 $\theta_i$ は、センサjとセンサj'とを結ぶ直線と、センサjとセンサj'との中心点と信号源iを結ぶ直線とのなす角度である。また、 $d_j$ はセンサjの位置を示すベクトルである。そして、例えば、各添字iと $\theta_i$ との対応が各周波数fにおいて同一となるように推定混合行列 $A(f)$ の列の並び替えを行い、パーミュテーション問題の解決を図る。

[0155] 2つ目の情報は、従来のICAを用いた方法と同様、分離信号成分の絶対値 $|Y_i(f, m)|$ の周波数間での相関である。すなわち、例えば、異なる周波数f1とf2において、同じ添字iに対する分離信号成分の絶対値の相関

[数46]

$$\text{cor}(v_i^{f1}, v_i^{f2}) = \frac{\langle v_i^{f1}(m) \cdot v_i^{f2}(m) \rangle_m}{\sqrt{\langle v_i^{f1}(m) \rangle_m} \cdot \sqrt{\langle v_i^{f2}(m) \rangle_m}} \dots (44)$$

(ただし、 $v_i^f(m) = |Y_i(f, m)| - \langle |Y_i(f, m)| \rangle_m$ 、 $\langle \cdot \rangle_m$ は時間mに関する「 $\cdot$ 」の平均値を示す。)

が最大化されるように推定混合行列 $A(f)$ の列の並び替えを行い、パーミュテーション問題の解決を図る。

なお、これらの処理に使用する分離信号は、分離信号生成部519の出力 $Y_1(f, m)$ , ...,  $Y_N(f, m)$ をフィードバックすることで得られる。

## [0156] [スケーリング問題解決部514での処理の詳細]

次に、スケーリング問題解決部514での処理の詳細を説明する。

パーミュテーション問題解決部513から推定混合行列 $A(f)$ を受け取ったスケーリング問題解決部514は、各列の大きさの任意性を解決するために、まず推定混合行列 $A(f)$ の各列(代表ベクトル) $a_i(f)$ に対し、正規化

$$a_i(f) \leftarrow a_i(f) / a_{ji}(f)$$

を行う(ステップS56)。なお、 $a_{ji}$ は、代表ベクトル $a_i(f)$ の $j$ 行目の要素である。また、 $j$ は各代表ベクトル $a_i(f)$ ごとに違うものを選んでもよいが、同じ $i$ に対してはすべての各周波数 $f$ において同じ $j$ を用いる必要がある。

## [0157] [分離信号生成処理の詳細]

次に、分離信号生成処理の詳細について説明する。

本形態の場合、センサの数が信号源の数に対して十分であるか否かによって分離信号の生成手順が異なる。

まず、センサの数が十分な場合( $M \geq N$ )は、簡単に分離信号を生成できる。すなわち、分離行列生成部518がスケーリング問題解決部514から推定混合行列 $A(f)$ を受け取り、そのムーア・ペンローズ型擬似逆行列 $A(f)^+$ ( $M=N$ の場合は逆行列 $A(f)^{-1}$ に一致)を分離行列 $W(f)$ として生成する(ステップS57)。生成された分離行列 $W(f)$ は、一時記憶部522に格納される。分離信号生成部519は、この分離行列 $W(f)$ と観測信号ベクトル $X(f, m)$ とを一時記憶部522から読み込み、これらを用い $Y(f, m) = W(f)X(f, m)$ の演算により、分離信号成分 $Y_1(f, m), \dots, Y_N(f, m)$ を生成する(ステップS58)。

[0158] 一方、センサの数が不十分な場合( $M < N$ )は、推定混合行列 $A(f)$ と観測信号ベクトル $X(f, m)$ とに対し、分離信号 $Y(f, m)$ は一意には定まらない。

[数47]

$$X(f, m) = A(f)Y(f, m) = \sum_{i=1}^N a_i(f)Y_i(f, m) \quad \dots(45)$$

を満たす $Y(f, m)$ が無数に存在するからである。源信号がスパース性を持つことに着目すると、無数の解のうち $L_1$ ノルム:



[数48]

$$L_1(Y(f, m)) = \sum_{i=1}^N |Y_i(f, m)| \quad \dots (46)$$

を最小にする解 $Y(f, m)$ が最も正確な分離信号成分となることが知られている(甘利俊一,「総論—人と機械はどのように見分け、聞き分けるのか—」, 電子情報通信学会誌, VOL. 87, No. 3, pp. 167, 2004年3月)。このような最小化基準で分離を行う場合は、分離のための行列 $W(f, m)$ が時変となり、分離行列生成部518は、時間 $m$ ごとに観測信号ベクトル $X(f, m)$ と推定混合行列 $A(f)$ とから時間依存の分離行列 $W(f, m)$ を算出し(ステップS57)、分離信号生成部519が、 $Y(f, m) = W(f, m)X(f, m)$ として分離信号成分 $Y_1(f, m), \dots, Y_N(f, m)$ を計算する(ステップS58)。

- [0159] しかし、 $L_1(Y(f, m))$ の最小化を厳密に行うのは計算量が大きいため、本形態では近似解法を用いて分離行列 $W(f, m)$ の生成を行う。この解法は、観測信号ベクトル $X(f, m)$ (或いはある時点での残差ベクトル $e$ )に最も方向が近い推定混合行列 $A(f)$ の列(代表ベクトル) $a_i(f)$ を順次選択していき、それらが $M$ 個選択されるまで繰り返すというものである。

図17は、本形態の近似解法を説明するためのフローチャートである。以下、このフローチャートに沿って近似解法を用いて分離行列 $W(f, m)$ を算出する処理を説明する。

- [0160] まず、列選択部516が、一時記憶部522から推定混合行列 $A(f)$ 及び観測信号ベクトル $X(f, m)$ を読み込み(ステップS61)、残差ベクトル $e$ を観測信号ベクトル $X(f, m)$ で初期化し、変数 $k$ に1を代入し(ステップS62)、これらの情報を一時記憶部522に格納する。

次に、列選択部516は、一時記憶部522内の変数 $k$ を参照し、 $k \leq M$ であるか否かを判断する(ステップS63)。ここで、 $k \leq M$ である場合、列選択部516は、

$$q(k) = \operatorname{argmax}_i |a_i(f)^H \cdot e| / \|a_i(f)\| \quad \dots (47)$$

となる $q(k)$ を選択し、その選択結果を一時記憶部522に格納する(ステップS64)。ここで、式(47)は、長さが正規化された列 $|a_i(f)^H| / \|a_i(f)\|$ と残差ベクトル $e$ の内積の絶対値を最大化するもの、すなわち残差ベクトル $e$ に最も方向が近い代表ベクトル

$a_i(f)$ を選択する演算を示している。残差ベクトル $e$ に方向が最も近い代表ベクトル $a_i(f)$ を選択する理由は、次の繰り返しでの残差ベクトル $e$ がより小さくなることで、以降の各 $Y_i(f, m)$ が小さくなり、最終的に式(46)で定義される $Y(f, m)$ の $L_1$ ノルムも小さくなると期待できるからである。

- [0161] 次に、列選択部516は、一時記憶部522に格納されている選択済みのすべての代表ベクトル $a_{q(1)}(f), \dots, a_{q(k)}(f)$ によって張られる部分空間を示す行列 $Q = [a_{q(1)}(f), \dots, a_{q(k)}(f)]$ を設定し(ステップS65)、 $P = Q(Q^H Q)^{-1} Q^H$ を算出する(ステップS66)。そして、列選択部516は、

$$e = X(f, m) - P \cdot X(f, m)$$

の演算によって残差ベクトル $e$ を更新して一時記憶部522に格納する(ステップS67)。

- [0162] ここで、 $P \cdot X(f, m)$ は、観測信号ベクトル $X(f, m)$ を部分空間 $Q$ に射影したもの、すなわち観測信号ベクトル $X(f, m)$ のうち、これまで選択された代表ベクトル $a_{q(1)}(f), \dots, a_{q(k)}(f)$ の線形和によって実現されるものである。残りの $e = X(f, m) - P \cdot X(f, m)$ は、他のベクトルによって実現され、具体的には以降のループ処理で選択される列(代表ベクトル) $a_{q(i)}$ によって実現される。

その後、順次、次の列を選択するため、列選択部516は、一時記憶部522の変数 $k$ に1を加えて新たな $k$ とし、ステップS63に戻る(ステップS68)。なお、残差ベクトル $e$ には、これまでに選択された代表ベクトル $a_{q(i)}$ と直交する成分しか含まれていないため、既に選択された代表ベクトルが、内積の絶対値 $|a_i(f)^H \cdot e| / \|a_i(f)\|$ の最大化基準のもと(ステップS64)で再び選択されることはない。

- [0163] そして、ステップS63で、列選択部516が $k \leq M$ と判断すると( $\min(M, N)$ 個の代表ベクトル $a_i(f)$ を選択したことに相当)、列選択部516はステップS64～68のループ処理を終了させる。この時点では、選択済みの $M$ 個の代表ベクトル $a_{q(i)}$ が全空間を張ることになるため、残差ベクトル $e$ は0となる。ステップS64～68のループ処理が終了すると、行列生成部517は、これまで選択された $M$ 個の代表ベクトル $a_{q(i)}$ を一時記憶部522から読み込み、ステップS63～68の処理で選択されなかった推定混合行列 $A(f)$ の $N-M$ 個の代表ベクトル(列ベクトル) $a_i(f)$ を0とした、

[数49]

$$a_i'(f, m) = \begin{cases} a_i(f) & i \in \{q(1), \dots, q(M)\} \\ 0 & i \notin \{q(1), \dots, q(M)\} \end{cases} \quad \dots (48)$$

という列ベクトル  $a_i'(f, m)$  を生成する(ステップS69)。さらに、行列生成部517は、式(48)の列ベクトル  $a_i'(f, m)$  を列とする行列  $A'(f, m) = [a_1'(f, m), \dots, a_N'(f, m)]$  (「選択された  $\min(M, N)$  個の代表ベクトル  $a_i(f)$  と  $\max(N-M, 0)$  個の0ベクトルとを列とした行列  $A'(f, m)$ 」に相当)を算出し、一時記憶部522に格納する(ステップS70)。なお、このように算出された行列  $A'(f, m)$  は、 $N \times M$  行列であるが、そのうち  $N-M$  個の行は0ベクトルである。

[0164] 分離行列生成部518は、このような行列  $A'(f, m)$  を一時記憶部522から読み出し、そのムーア・ペンローズ型擬似逆行列  $A'(f, m)^+$  を分離行列  $W(f, m)$  として生成する(ステップS71)。これは、 $N$  個の代表ベクトル  $a_i(f)$  の中の0個以上の当該代表ベクトルを0ベクトルに置換した  $M$  行  $N$  列の行列のムーア・ペンローズ型擬似逆行列である  $N$  行  $M$  列の分離行列  $W(f, m)$  に相当する。

[0165] 生成された分離行列  $W(f, m)$  は、一時記憶部522に格納される。分離信号生成部519は、この分離行列  $W(f, m)$  と、観測信号ベクトル  $X(f, m)$  と一時記憶部522から読み込み、 $Y(f, m) = W(f, m)X(f, m)$  として分離信号成分  $Y_1(f, m), \dots, Y_N(f, m)$  を生成し、一時記憶部522に格納する(ステップS58)。なお、このように生成された分離信号成分  $Y_1(f, m), \dots, Y_N(f, m)$  のうち  $N-M$  個の要素は必ず0になる。すなわち、ある離散時間  $m$  のみについてステップS61〜S71の処理を行っただけでは、最大  $M$  個の分離信号成分しか知ることができない。そのため、本形態では、これまで説明した  $M$  個の代表ベクトル  $a_i(f)$  の選択、行列  $A'(f, m)$  の生成、分離行列  $W(f, m)$  の算出、分離信号ベクトル  $Y(f, m)$  の算出、及び時間領域の信号値  $y_1(t), \dots, y_N(t)$  への変換の処理を、離散時間  $m$  ごとに行う。これにより、すべての分離信号成分を知ることができる。

[0166] <本形態の効果>

[ $N > M$ でのブラインド信号分離]

以上説明した通り本形態では、センサ数が少ない( $N > M$ )状況でも、源信号がス

パース性を備えていればブラインド信号分離が達成できる。その結果、センサの数を低減でき装置のコスト低減にも貢献できる。

[正規化の効果]

図18～23は、正規化部512aaで行われた正規化の効果を例示したプロットである。これらの例は、残響時間130msの部屋で1つ或いは2つの音声を2つのマイクで観測した場合の2773Hzにおける観測信号ベクトル $X(f, m)$ のプロットである。なお、これらは2個のマイクで観測した例であるが、観測信号ベクトル $X(f, m)$ は周波数領域における複素ベクトルであるため、実数では4次元空間でのベクトルとなる。そのため、図4～9ではその4次元を4つの2次元空間に射影して表示した。なお、これらの図の「imag」は各観測信号の虚数項を示し、「real」は実数項を示す。また、 $X_1$ は第1のマイクで観測された観測信号に係るデータを示し、 $X_2$ は第2のマイクで観測された観測信号に係るデータを示している。

[0167] まず、1音源の場合における正規化の効果を図18～20に示す。

図18は、正規化していない観測信号ベクトル $X(f, m)$ のプロットである。この例では、原点を中心にクラスタが形成されているが、そのクラスタから源信号1に関する代表ベクトル $a_1(f)$ について有益な情報は得られない。一方、図19は、式(36)により正規化した観測信号ベクトル $X(f, m)$ のプロットである。この例では、原点からある特定の方向にサンプルが散布されている。この方向が推定すべき代表ベクトル $a_1(f)$ に対応する。これは代表ベクトル $a_1(f)$ を決定するうえでの有益な情報となる。また、図20は、式(37)により正規化した観測信号ベクトル $X(f, m)$ のプロットである。この例では、原点から離れた箇所にクラスタが形成されている。このクラスタの中心と原点を結ぶベクトルが推定すべき代表ベクトル $a_1(f)$ に対応する。

[0168] 次に、2音源の場合における正規化の効果を図21～23に示す。

図21は、正規化していない観測信号ベクトル $X(f, m)$ のプロットである。この例の場合も、1音源の場合と同様に、2つの源信号に関して有益な情報は得られない。図22は、式(36)により正規化した観測信号ベクトル $X(f, m)$ のプロットである。この例の場合、原点から2つの方向にサンプルが散布している。そして、これら方向が推定すべき代表ベクトル $a_1(f)$ 、 $a_2(f)$ に対応する。図23は、式(37)により正規化した観測信号

ベクトル $X(f, m)$ のプロットである。この例では、原点から離れた箇所に2個のクラスタを形成されていることがわかる。そして、このクラスタの中心と原点を結ぶベクトルが推定すべき代表ベクトル $a_1(f)$ ,  $a_2(f)$ に対応する。

[0169] [近似解法を用いた分離行列生成の効果]

前述したように、 $N > M$ の場合の分離行列 $W(f, m)$ の生成において最小化を厳密に行った場合、その計算量は膨大なものになってしまう。例えば、 $N$ 個の代表ベクトル $a_1(f), \dots, a_N(f)$ から $M$ 個を選択する組合せは $C_{N, M}$ 個あるため、厳密に $L_1$ ノルム(式(46))を極小化する組合せをみつけようとすれば、 $C_{N, M}$ 個の組についての並び替えの処理等が必要となる。しかし、図17に示した近似解法では、センサの数 $M$ に相当する回数のループを繰り返せばよく、計算量は少なく済む。

[0170] なお、本形態では、センサの数が信号源の数に対して十分であるか否か、すなわち $N \leq M$ であるか否かによって、分離行列 $W(f, m)$ の生成手順を相違させることとした。しかし、センサの数が信号源の数に対して十分であるか否かに係わらず同じルーチンを用いて分離行列 $W(f, m)$ を生成することとしてもよい。

図24は、このような例を説明するためのフローチャートである。

この変形例の場合、 $N \leq M$ であるか否かに係わらず、まず、列選択部516が、一時記憶部522から推定混合行列 $A(f)$ 及び観測信号ベクトル $X(f, m)$ を読み込み(ステップS81)、残差ベクトル $e$ を観測信号ベクトル $X(f, m)$ で初期化し、変数 $k$ に1を代入する(ステップS82)。そして、列選択部516が、 $k \leq \min(M, N)$ であるか否かを判断し(ステップS83)、 $k \leq \min(M, N)$ であれば、 $|a_{q(u)}(f)^H \cdot e| / \sqrt{a_{q(u)}(f)^H \cdot a_{q(u)}(f)}$ は $\alpha$ の共役転置行列)を最大にする列 $a_{q(u)}(f)$ を選択し(ステップS84)、選択済みのすべての列 $a_{q(u)}(f)$  ( $u=1, \dots, k$ )によって張られる部分空間を示す行列 $Q = [a_{q(1)}(f), \dots, a_{q(k)}(f)]$ を設定し(ステップS85)、 $P = Q(Q^H Q)^{-1} Q^H$ を算出し(ステップS86)、 $X(f, m) - P \cdot X(f, m)$ の演算結果によって残差ベクトル $e$ を更新し(ステップS87)、変数 $k$ に1を加えた値を新たな $k$ として(ステップS88)、ステップS83に戻る。すなわち、ステップS83〜88の処理を $\min(M, N)$ 回繰り返す。なお、 $\min(M, N)$ とは、 $M$ 及び $N$ の何れか小さい方の値を意味し、 $\max(N-M, 0)$ とは、 $N-M$ 及び0の何れか大きい方の値を意味する。

[0171] その後、列選択部516は、これまで選択した $\min(M, N)$ 個の代表ベクトル $a_{q(i)}$ を一時記憶部522に格納する。

次に、行列生成部517は、一時記憶部522からこれら $\min(M, N)$ 個の代表ベクトル $a_{q(i)}$ を読み込み、

[数50]

$$a_i'(f, m) = \begin{cases} a_i(f) & i \in \{q(1), \dots, q(\min(M, N))\} \\ 0 & i \notin \{q(1), \dots, q(\min(M, N))\} \end{cases} \quad \dots (49)$$

という列ベクトル $a_i'(f, m)$ を生成し(ステップS89)、ステップS83〜88で選択された $\min(M, N)$ 個の代表ベクトル $a_i(f)$ と、 $\max(N-M, 0)$ 個の0ベクトルとを列とした行列 $A'(f, m) = [a_1'(f, m), \dots, a_N'(f, m)]$ を生成する(ステップS90)。このように生成された行列 $A'(f, m)$ は一時記憶部522に格納された後、分離行列生成部518に読み込まれ、分離行列生成部518は、そのムーア・ペンローズ型擬似逆行列 $A(f, m)^+$ ( $M=N$ の場合は逆行列 $W^{-1}$ に一致)を分離行列 $W(f, m)$ として生成する(ステップS91)。なお、これは $N$ 個の上記代表ベクトル $a_i(f)$ の中の0個以上の当該代表ベクトルを0ベクトルに置換した $M$ 行 $N$ 列の行列のムーア・ペンローズ型擬似逆行列である $N$ 行 $M$ 列の分離行列 $W(f, m)$ に相当する。

[0172] [変形例等]

なお、本発明は上述の各実施の形態に限定されるものではない。例えば、第1の実施の形態から第8の実施の形態では、抽出信号を時間領域に戻してから統合を行うこととしていたが、バイナリマスクを用いる場合には、周波数領域で信号統合を行ってから時間領域に変換してもよい。

図25は、周波数領域で信号統合を行ってから時間領域に変換する際の構成を例示したブロック図の一部である。この図の構成は、図1における限定信号分離部60-k、時間領域変換部70-k及び信号統合部80の代わりに設けられる構成である。

[0173] この例では、すべての系列 $l$ の限定信号分離部601-kから出力された周波数領域の信号値 $Y_{kq}^{\Pi kq}(f, m)$ について、信号統合部602が周波数領域にて信号の統合を行った後に、時間領域変換部603が時間領域へ変換する。ここで、信号統合部602は、ある周波数 $f$ において、同じタグ $a_i$ を持つ分離信号 $Y_{kq}^{\Pi kq}(f, m)$ が一つしかない

場合、

$$Y_i(f, m) = Y_{kq}^{\Pi_{kq}}(f, m)$$

として分離信号値を求める。また、ある周波数 $f$ において、同じタグ $a_i$ を持つ分離信号 $Y_{kq}^{\Pi_{kq}}(f, m)$ が二つ以上ある場合、 $Y_i(f, m)$ は、例えば同じタグ $a_i$ を持つ分離信号 $Y_{kq}^{\Pi_{kq}}(f, m)$ の平均として

[数51]

$$Y_i(f, m) = \frac{1}{K} \sum_{\Pi_{kq}=a_i} Y_{kq}^{\Pi_{kq}}(f, m)$$

( $K$ は同じタグ $a_i$ を持つ分離信号の個数)

として求められる。

[0174] そして最後に、時間領域変換部603が、例えば短時間逆フーリエ変換などにより、周波数領域で統合された出力信号値 $Y_i(f, m)$ を時間領域 $y_i(t)$ に変換する。

また、第1の実施の形態から第8の実施の形態では、各分離信号にタグを付与して信号の統合処理を行うこととしたが、各分離信号にタグを付けるのではなく、各系統 $k$ において分離される信号に対応する $V$ 個の代表値の集合 $G_k$ を一時記憶部90に保持しておき、出力信号の統合を行うこととしてもよい。

[0175] 具体的には、例えば、複数の系統において $G_k$ が同じ代表値を含むことがない場合は、すべての分離信号 $y_{kq}(t)$ を最終的な分離信号 $y_i(t)$  ( $i=1, \dots, N$ )として出力する。また周波数領域におけるすべての分離信号 $Y_{kq}(f, m)$ を周波数領域における最終的な分離信号 $Y_i(f, m)$  ( $i=1, \dots, N$ )とし、時間領域の信号に変換しても良い。

[0176] また、複数の系統において $G_k$ が同じ代表値を $K$ 個 ( $K \geq 2$ ) 含む場合は、 $k$ 系統の分離信号 $y_{kq}(t)$  ( $q=1, \dots, V_k$ 、 $V_k/V_k$ は $G_k$ の個数)と $k'$ 系統の分離信号 $y_{k'r}(t)$  ( $r=1, \dots, V_{k'}$ )についてのすべての組合せで信号の相関を計算し、相関の高いもの $K$ 個について $y_{kq}(t)$ と $y_{k'r}(t)$ の平均をとる。これを、同じ代表値を含む複数の系統について繰り返し、信号を統合する。また、同じ操作を周波数領域におけるすべての分離信号について行うことで、周波数領域で信号を統合し、その後で時間領域の信号に変換してもよい。

[0177] さらに、上述の第1から第9の各実施の形態を複合したシステムによって信号分離

を行ってもよい。

例えば、[第8の実施の形態]の方法で代表ベクトルを求め、その後[第2の実施の形態]の方法で限定信号を分離することとしてもよい。具体的には、例えば、代表値計算部430(図13)で求めた代表ベクトルを用い、[第2の実施の形態]における(18)式のかわりに

[数52]

$$M_k(f, m) = \begin{cases} 1 & D(X(f, m), a_k(f)) < \min_{k \neq j} D(X(f, m), a_j(f)) \\ 0 & \text{otherwise} \end{cases}$$

として  $M_k(f, m)$  を、

$$M_{DC}(f, m) = \begin{cases} 1 & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

として  $M_{DC}(f, m)$  を求め ( $M_k(f, m)$ ,  $M_{DC}(f, m)$  は、図8参照。)、あとは[第2の実施の形態]の限定信号分離部160-kと同じ手順で限定信号を分離する。

ここで、上述の  $M_k(f, m)$ ,  $M_{DC}(f, m)$  を求めず、直接

[数53]

$$\hat{X}_k(f, m) = \begin{cases} X(f, m) & D(X(f, m), a_k(f)) < \min_{k \neq j} D(X(f, m), a_j(f)) \\ 0 & \text{otherwise} \end{cases}$$

にて  $\hat{X}_k(f, m) = M_k X$  (マスク作成部151-k及び積演算部161-k

(図8)の処理に対応)を求め、

$$\hat{X}_k(f, m) = \begin{cases} X(f, m) & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

にて限定信号値を生成してもよい(マスク作成部151-k及び限定信号抽出部152-k(図8)の処理に対応)。

[0178] また、マスク作成部151-k(図8)にて  $M_k(f, m)$  を作成せず、かわりに、混合過程推定部162-k(図8)にて、代表値計算部430(図13)にて求めた代表ベクトル  $[a_1, \dots, a_N]$  ( $a_1$  は縦ベクトル) をまとめて  $H$  とし、これを推定混合行列としてもよい。

また、上述の各実施の形態ではフーリエ変換／逆フーリエ変換によって時間領域一周波数領域間の変換を行うこととしたが、wavelet変換、DFTフィルタバンク、ポリフェ



イズフィルタバンクなどを用い、この変換を行うこととしてもよい(例えば、「R. E. Crochiere, L. R. Rabiner, "Multirate Digital Signal Processing." Eaglewood Cliffs, NJ: Printice-Hall, 1983 (ISBN 0-13-605162-6) )。

- [0179] また、上述の第1から第9の各実施の形態をコンピュータで構成する場合は以下のようになる。

図26は、各実施の形態をコンピュータで構成した信号分離装置610の例である。

この例の信号分離装置610は、CPU(Central processing Unit: 中央処理装置) 620、RAM(Random Access Memory) 630、ROM(Read Only Memory) 640、外部記憶装置650、入力部660、インタフェース670及びバス680を有している。

- [0180] CPU620は、例えば、演算部621、制御部622及びレジスタ623を有するCISC(Complex Instruction Set Computer)方式、RISC(Reduced Instruction Set Computer)方式等の中央処理装置である。また、レジスタ623は、例えばDRAM(Dynamic Random Access Memory)、SRAM(Static Random Access Memory)等の動作が高速なメモリである。

また、RAM630は、例えば、DRAM、SRAM、フラッシュメモリ、NV(Nonvolatile) RAM等の読書き可能な半導体メモリである。またROM640は、例えば、MROM(Mask Read Only Memory)等の読み出し専用の半導体メモリであり、各種プログラムやデータ等が記憶されている。

- [0181] 外部記憶装置650は、例えば、ハードディスク装置、フレキシブルディスク、磁気テープ等の磁気記録装置、DVD-RAM(Random Access Memory)、CD-R(Recordable)／RW(ReWritable)等の光ディスク装置、MO(Magneto-Optical disc)等の光磁気記録装置、EEP-ROM(Electronically Erasable and Programmable-Read Only Memory)、フラッシュメモリ(flash memory)等の半導体メモリ等である。

また、入力部660は、例えば、キーボード、マウス、ジョイスティック等の入力デバイスである。また、インタフェースは、例えば、データの入力、出力、或いはその双方を行う入力／出力ポートであり、例えば、センサ、通信ボード、記憶装置等の各種装置が接続可能となっている。

[0182] さらに、バス680は、例えば、データバス、アドレスバス、コントロールバス等によって構成され、CPU620、RAM630、ROM640、外部記憶装置650、入力部660及びインタフェース670でのデータのやり取りが可能なようにこれらを電氣的に接続する。

また、信号分離装置610における処理の内容は、例えば、信号分離プログラムに記述され、このような信号分離プログラムは、例えば、コンピュータで読み取り可能な記録媒体に記録される。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよいが、具体的には、例えば、磁気記録装置として、ハードディスク装置、フレキシブルディスク、磁気テープ等を、光ディスクとして、DVD (Digital Versatile Disc)、DVD-RAM (Random Access Memory)、CD-ROM (Compact Disc Read Only Memory)、CD-R (Recordable) / RW (ReWritable) 等を、光磁気記録媒体として、MO (Magneto-Optical disc) 等を、半導体メモリとしてEEP-ROM (Electrically Erasable and Programmable-Read Only Memory) 等を用いることができる。

[0183] また、この信号分離プログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

信号分離装置610において処理を実行する場合、例えばまず、可搬型記録媒体に記録された信号分離プログラムもしくはサーバコンピュータから転送された信号分離プログラムを、外部記憶装置650のプログラム領域651にダウンロードする。

[0184] また、各センサにおいて観測された時間領域の各観測信号 $x_j(t)$  ( $j=1, \dots, M$ )も、事前に外部記憶装置650のデータ領域652に格納される。この各観測信号 $x_j(t)$ の格納は、センサから送られた各観測信号 $x_j(t)$ をインタフェース670に入力し、バス680を通じて外部記憶装置650に格納することとしてもよく、事前に別の装置で各観測信号 $x_j(t)$ を外部記憶装置650に格納しておき、この外部記憶装置650をバス680に接続する構成としてもよい。

次に、例えば、CPU620の制御部622の制御のもと、外部記憶装置650のプログ

ラム領域651から信号分離プログラムが順次読み出され、RAM630のプログラム領域631に格納される。RAM630に格納された信号分離プログラムは、CPU620に読み込まれ、CPU620の制御部622は、この信号分離プログラムの内容に従い、データの入出力、演算部621での演算、レジスタ623へのデータ格納等の各処理を実行する。

[0185] CPU620による処理が開始されると、CPU620は、例えば外部記憶装置650のデータ領域652の各観測信号 $x_j(t)$ を読み出し、例えばRAM630のデータ領域632に書き込む。その後CPU620は、制御部622の制御のもと、RAM630のプログラム領域631の信号分離プログラム、及びデータ領域632の信号分離プログラムを順次抽出しつつ、上述の各処理を実行する。なお、例えばRAM630或いは外部記憶装置650が、第1から第9の実施の形態における記憶部2, 501として機能し、RAM630或いはレジスタ623が、第1から第9の実施の形態における一次記憶部90, 522として機能する。

[0186] また、このプログラムの別の実行形態として、CPU620が可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このCPU620にサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。

[0187] さらに、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

#### 産業上の利用可能性

[0188] 本発明により、例えば、様々なノイズ・妨害信号が存在する環境下においても、目的信号を精度よく分離抽出することが可能となる。例えば、オーディオ分野に応用した場合、音声認識機の入力マイクロホンと話者が離れた位置にあり、マイクロホンが

目的話者音声以外の音まで集音してしまうような状況でも、目的音声を分離抽出することにより、認識率の高い音声認識系を構築することができる。

### 請求の範囲

- [1]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離方法であって、
- 前記センサにおいて観測された観測信号値を周波数領域の信号値に変換する手順と、
- 前記周波数領域の信号値を用い、前記センサ間における観測値の相対値(相対値の写像も含む)を、各周波数において算出する手順と、
- 前記相対値を $N$  個のクラスにクラスタリングする手順と、
- 前記の各クラスの代表値を算出する手順と、
- 前記代表値を用い、前記周波数領域の信号値から、 $V$  ( $2 \leq V \leq M$ ) 個の信号源から発せられた信号から成る混合信号の値を抽出するためのマスクを作成する手順と、
- 前記マスクを用い、前記周波数領域の信号値から前記混合信号の値を抽出する手順と、
- 前記混合信号の値から $V$  個の信号の値を分離抽出する手順と、
- を有することを特徴とする信号分離方法。
- [2] 請求項1記載の信号分離方法であって、
- 前記マスクは、
- $V$  個の前記代表値を含む所定の範囲内にある前記相対値に対してハイレベル値をとり、前記所定の範囲内にない前記代表値に対してローレベル値をとる関数であり、
- 前記マスクを用い、前記周波数領域の信号値から前記混合信号の値を抽出する手順は、
- 前記周波数領域の信号値に前記マスクを乗じる手順である、
- ことを特徴とする信号分離方法。
- [3] 請求項1記載の信号分離方法であって、
- 前記マスクは、
- $V$  個の前記代表値を含む所定の範囲内にある前記相対値に対してローレベル値をとり、前記所定の範囲内にない前記代表値に対してハイレベル値をとる関数であり、
- 前記マスクを用い、前記周波数領域の信号値から前記混合信号の値を抽出する手

順は、

前記周波数領域の信号値に前記マスクを乗じた値を、前記周波数領域の信号の値から減算する手順である、

ことを特徴とする信号分離方法。

[4] 請求項2記載の信号分離方法であって、

前記マスクは、

前記相対値の変化に伴う前記ハイレベル値から前記ローレベル値への推移が連続的な関数である、

ことを特徴とする信号分離方法。

[5] 請求項1記載の信号分離方法であって、

前記代表値を用い、前記周波数領域の信号値から、 $V$  ( $2 \leq V \leq M$ ) 個の信号源から発せられた信号から成る混合信号の値を抽出するためのマスクを作成する手順は

、  
死角型ビームフォーマ(NBF)の指向特性を利用して前記マスクを作成する手順である、

ことを特徴とする信号分離装置。

[6] 請求項1記載の信号分離方法であって、

前記代表値を用い、前記周波数領域の信号値から、 $V$  ( $2 \leq V \leq M$ ) 個の信号源から発せられた信号から成る混合信号の値を抽出するためのマスクを作成する手順は

、  
 $f$ を周波数の変数とし、 $\theta_1$ を $V$ 個の前記代表値に対応する信号源の推定方向の何れか1つとし、 $\theta_i$  ( $i=2, \dots, N-V+1$ )を前記 $V$ 個の代表値以外の前記代表値に対応する各信号源の推定方向とし、 $j=1, \dots, N-V+1$ とし、 $d_j$ をセンサ1とセンサ $j$ との距離とし、 $v$ を信号の速さとし、 $\tau_{ji} = (d_j/v) \cos \theta_i$ とした場合における、 $ji$ 要素が $\exp(j2\pi f \tau_{ji})$ である $(N-V+1) \times (N-V+1)$ の遅延行列 $H_{NBF}(f)$ を生成する手順と、  
遅延行列 $H_{NBF}(f)$ の逆行列 $W(f) = H_{NBF}^{-1}(f)$ をNBF行列 $W(f)$ として算出する手順と、

前記NBF行列 $W(f)$ の1行目の要素を $W_{1k}(f)$ とし、 $\theta$ を信号の到来方向の変数と

した場合における、指向特性関数

[数54]

$$F(f, \theta) = \sum_{k=1}^{N-V+1} W_{1k}(f) \exp(j2\pi f d_k \cos \theta / v)$$

を生成する手順と、

前記指向特性関数 $F(f, \theta)$ を用いて前記マスクを生成する手順と、

を具備することを特徴とする信号分離方法。

[7] 請求項1記載の信号分離方法であって、

前記代表値を用い、前記周波数領域の信号値から、 $V(2 \leq V \leq M)$ 個の信号源から発せられた信号から成る混合信号の値を抽出するためのマスクを作成する手順は

、

$V$ 個の前記代表値を含む所定の範囲内にある前記相対値に対してハイレベル値をとり、前記所定の範囲内にない前記代表値に対してローレベル値をとり、相対値の変化に伴う当該ハイレベル値から当該ローレベル値への推移が不連続な関数であるバイナリマスクに単峰性関数を畳み込んだ関数を生成する手順と、

前記バイナリマスクに単峰性関数を畳み込んだ関数に前記相対値を代入した関数を前記マスクとして生成する手順と、

を具備することを特徴とする信号分離方法。

[8] 請求項1記載の信号分離方法であって、

前記代表値を用い、前記周波数領域の信号値から、 $V(2 \leq V \leq M)$ 個の信号源から発せられた信号から成る混合信号の値を抽出するためのマスクを作成する手順は

、

前記相対値が $V$ 個の前記代表値を含む所定の範囲内の下限値 $a_{\min}$ である場合に0をとる第1の奇関数と、前記相対値が前記所定の範囲内の上限値 $a_{\max}$ である場合に0をとる第2の奇関数との差の写像から得られる単峰性の関数を、前記マスクとして生成する手順である、

ことを特徴とする信号分離方法。

[9] 請求項2或いは3に記載の信号分離方法であって、

前記マスクは、

前記ハイレベル値から前記ローレベル値への推移が不連続な関数である、  
ことを特徴とする信号分離方法。

- [10]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離方法であって、

前記センサにおいて観測された観測信号値を周波数領域の信号値に変換する手順と、

前記周波数領域の信号値を用い、前記センサ間における観測値の相対値(相対値の写像も含む)を各周波数において算出する手順と、

前記相対値を $N$  個のクラスにクラスタリングする手順と、

前記の各クラスの代表値を算出する手順と、

1 個の前記代表値を含む所定の範囲内にある前記相対値に対してハイレベル値をとり、前記所定の範囲内にない前記代表値に対してローレベル値をとり、前記相対値の変化に伴う当該ハイレベル値から当該ローレベル値への推移が連続的な関数であるマスクを作成する手順と、

前記周波数領域の信号値に前記マスクを乗じ、1 個の信号源から発せられた信号の値を抽出する手順と、

を有することを特徴とする信号分離方法。

- [11]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離方法であって、

前記センサにおいて観測された観測信号値  $x_1(t), \dots, x_M(t)$  を周波数領域の信号値  $X_1(f, m), \dots, X_M(f, m)$  に変換する手順と、

前記周波数領域の信号値  $X_1(f, m), \dots, X_M(f, m)$  からなる第1のベクトル  $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$  を、周波数  $f$  ごとに  $N$  個ずつのクラス  $C_i(f)$  ( $i = 1, \dots, N$ ) にクラスタリングする手順と、

前記各クラス  $C_i(f)$  を代表する第2のベクトル  $a_i(f)$  を算出する手順と、

前記第2のベクトル  $a_i(f)$  から  $V$  ( $1 \leq V \leq M$ ) 個の第3のベクトル  $a_p(f)$  ( $p = 1, \dots, V$ ) を抽出する手順と、



前記第3のベクトル $a_p(f)$ の集合を $G_k$ とし、 $G_k^c$ を $G_k$ の補集合とし、 $D(\alpha, \beta)$ をベクトル $\alpha$ と $\beta$ とのマハラノビス平方距離とした場合における、

[数55]

$$M(f, m) = \begin{cases} 1 & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

で示されるマスク $M(f, m)$ を生成する手順と、

前記マスク $M(f, m)$ と前記第1のベクトル $X(f, m)$ との積を演算し、 $V$ 個の前記信号源から発せられた信号の値を抽出する手順と、

を有することを特徴とする信号分離方法。

[12]  $N(N \geq 2)$ 個の信号が混合し、 $M$ 個のセンサで観測された状況において信号の分離抽出を行う信号分離方法であって、

前記センサにおいて観測された観測信号値 $x_1(t), \dots, x_M(t)$ を周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ に変換する手順と、

前記周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ からなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$ を、周波数 $f$ ごとに $N$ 個ずつのクラスタ $C_i(f)$  ( $i = 1, \dots, N$ )にクラスタリングする手順と、

前記各クラスタ $C_i(f)$ を代表する第2のベクトル $a_i(f)$ を算出する手順と、

前記第2のベクトル $a_i(f)$ から $V(1 \leq V \leq M)$ 個の第3のベクトル $a_p(f)$  ( $p = 1, \dots, V$ )を抽出する手順と、

前記第1のベクトル $X(f, m)$ に対し、前記第3のベクトル $a_p(f)$ の集合を $G_k$ とし、 $G_k^c$ を $G_k$ の補集合とし、 $D(\alpha, \beta)$ をベクトル $\alpha$ と $\beta$ とのマハラノビス平方距離とした場合における、

[数56]

$$\max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f))$$

を満たすか否かを判断し、満たすと判断した前記第1のベクトル $X(f, m)$ を、 $V$ 個の前記信号源から発せられた信号の値として抽出する手順と、

を有することを特徴とする信号分離方法。

- [13] 請求項11或いは12記載の信号分離方法であって、  
前記クラスタリングする手順は、

[数57]

$$\text{sign}(X_j(f, m)) \leftarrow \begin{cases} X_j(f, m) / |X_j(f, m)| & (|X_j(f, m)| \neq 0) \\ 0 & (|X_j(f, m)| = 0) \end{cases}$$

を算出し、

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \text{sign}(X_j(f, m)) & (|X_j(f, m)| \neq 0) \\ X(f, m) & (|X_j(f, m)| = 0) \end{cases}$$

の演算を行った後に行われる、

ことを特徴とする信号分離方法。

- [14] 請求項13記載の信号分離方法であって、  
前記クラスタリングする手順は、前記の

[数58]

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \text{sign}(X_j(f, m)) & (|X_j(f, m)| \neq 0) \\ X(f, m) & (|X_j(f, m)| = 0) \end{cases}$$

を演算を行った後、さらに

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \|X(f, m)\| & (\|X(f, m)\| \neq 0) \\ X(f, m) & (\|X(f, m)\| = 0) \end{cases}$$

(ただし、 $\|X(f, m)\|$ は $X(f, m)$ のノルム)

の演算を行った後に行われる、

ことを特徴とする信号分離方法。

- [15]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離方法であって、

前記センサにおいて観測された観測信号値  $x_1(t), \dots, x_M(t)$  を、周波数領域の信号値  $X_1(f, m), \dots, X_M(f, m)$  に変換する手順と、

前記周波数領域の信号値  $X_1(f, m), \dots, X_M(f, m)$  からなる第1のベクトル  $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]^T$  を、周波数  $f$  ごとに  $N$  個ずつのクラス  $C_i(f)$  ( $i =$

1, ..., N)にクラスタリングする手順と、

前記各クラスタ $C_i(f)$ を代表する第2のベクトル $a_i(f)$ を算出する手順と、

N個の前記第2のベクトル $a_i(f)$ の中の0個以上の当該第2のベクトルを0ベクトルに置換したM行N列の行列のムーア・ペンローズ型擬似逆行列であるN行M列の分離行列 $W(f, m)$ を算出する手順と、

$Y(f, m) = W(f, m)X(f, m)$ の演算により、分離信号ベクトル $Y(f, m) = [Y_1(f, m), \dots, Y_N(f, m)]^T$ を算出する手順と、

を有することを特徴とする信号分離方法。

[16] 請求項15記載の信号分離方法であって、

前記分離行列 $W(f, m)$ を算出する手順は、

$\min(M, N)$ 個の前記第2のベクトル $a_i(f)$ を選択し、選択した $\min(M, N)$ 個の前記第2のベクトル $a_i(f)$ と $\max(N-M, 0)$ 個の0ベクトルとを列とした行列 $A'(f, m)$ を生成し、前記行列 $A'(f, m)$ のムーア・ペンローズ型擬似逆行列を前記分離行列 $W(f, m)$ として算出する手順である、

ことを特徴とする信号分離方法。

[17] 請求項15記載の信号分離方法であって、

$N > M$ である場合における前記分離行列 $W(f, m)$ を算出する手順は、

離散時間 $m$ ごとに、M個の前記第2のベクトル $a_i(f)$ を選択し、前記選択したM個の前記第2のベクトル $a_i(f)$ と $N-M$ 個の0ベクトルとを列とした行列 $A'(f, m)$ を生成し、前記行列 $A'(f, m)$ のムーア・ペンローズ型擬似逆行列を、時間依存の前記分離行列 $W(f, m)$ として算出する手順であり、

$N \leq M$ である場合における前記分離行列 $W(f, m)$ を算出する手順は、

前記各クラスタ $C_i(f)$ のN個の前記第2のベクトル $a_i(f)$ からなる行列のムーア・ペンローズ型擬似逆行列を、時不変の前記分離行列 $W(f, m)$ として算出する手順である、

ことを特徴とする信号分離方法。

[18] 請求項15記載の信号分離方法であって、

前記クラスタリングする手順は、

[数59]

$$\text{sign}(X_j(f, m)) \leftarrow \begin{cases} X_j(f, m) / |X_j(f, m)| & (|X_j(f, m)| \neq 0) \\ 0 & (|X_j(f, m)| = 0) \end{cases}$$

を算出し、

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \text{sign}(X_j(f, m)) & (|X_j(f, m)| \neq 0) \\ X(f, m) & (|X_j(f, m)| = 0) \end{cases}$$

の演算を行った後に行われる、

ことを特徴とする信号分離方法。

- [19] 請求項18記載の信号分離方法であって、  
前記クラスタリングする手順は、前記の

[数60]

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \text{sign}(X_j(f, m)) & (|X_j(f, m)| \neq 0) \\ X(f, m) & (|X_j(f, m)| = 0) \end{cases}$$

を演算を行った後、さらに

$$X(f, m) \leftarrow \begin{cases} X(f, m) / \|X(f, m)\| & (\|X(f, m)\| \neq 0) \\ X(f, m) & (\|X(f, m)\| = 0) \end{cases}$$

(ただし、 $\|X(f, m)\|$ は $X(f, m)$ のノルム)

の演算を行った後に行われる、

ことを特徴とする信号分離方法。

- [20] 請求項16記載の信号分離方法であって、

前記 $\min(M, N)$ 個の前記第2のベクトル $a_i(f)$ を選択する手順は、

第4のベクトル $e$ を前記第1のベクトル $X(f, m)$ で初期化した後、 $a_{q(u)}(f) / \|a_{q(u)}(f)\|$ と前記第4のベクトル $e$ の内積の絶対値を最大化する前記第2のベクトル $a_{q(u)}(f)$ を選択し、選択済みのすべての前記第2のベクトル $a_{q(u)}(u=1, \dots, k)$ によって張られる部分空間を示す行列 $Q = [a_{q(1)}(f), \dots, a_{q(k)}(f)]$ を設定し、 $P = Q(Q^H Q)^{-1} Q^H$ を算出し、 $e = X(f, m) - P \cdot X(f, m)$ の演算結果によって第4のベクトル $e$ を更新する処理を $\min(M, N)$ 回繰り返す手順である、

ことを特徴とする信号分離方法。

- [21]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離装置であって、

前記センサにおいて観測された観測信号値を格納する記憶部と、

前記記憶部に接続され、

前記観測信号値を周波数領域の信号値に変換し、

前記周波数領域の信号値を用い、前記センサ間における観測値の相対値(相対値の写像も含む)を、各周波数において算出し、

前記相対値を $N$  個のクラスにクラスタリングし、

前記の各クラスの代表値を算出し、

前記代表値を用い、前記周波数領域の信号値から、 $V$  ( $2 \leq V \leq M$ ) 個の信号源から発せられた信号から成る混合信号の値を抽出するためのマスクを作成し、

前記マスクを用い、前記周波数領域の信号値から前記混合信号の値を抽出し、

前記混合信号の値から $V$  個の信号の値を分離抽出する、処理を実行するプロセスと、

を有することを特徴とする信号分離装置。

- [22]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離装置であって、

前記センサにおいて観測された観測信号値を格納する記憶部と、

前記記憶部に接続され、

前記観測信号値を周波数領域の信号値に変換し、

前記周波数領域の信号値を用い、センサ間における観測値の相対値(相対値の写像も含む)を各周波数において算出し、

前記相対値を $N$  個のクラスにクラスタリングし、

前記の各クラスの代表値を算出し、

1 個の前記代表値を含む所定の範囲内にある前記相対値に対してハイレベル値をとり、前記所定の範囲内でない前記代表値に対してローレベル値をとり、前記相対値の変化に伴う当該ハイレベル値から当該ローレベル値への推移が連続的な関数で

あるマスクを作成し、

前記周波数領域の信号値に前記マスクを乗じ、1個の信号源から発せられた信号の値を抽出する、処理を実行するプロセッサと、

を有することを特徴とする信号分離装置。

- [23]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離装置であって、

前記センサにおいて観測された観測信号値  $x_1(t), \dots, x_M(t)$  を格納する記憶部と、

前記記憶部に接続され、

前記観測信号値  $x_1(t), \dots, x_M(t)$  を周波数領域の信号値  $X_1(f, m), \dots, X_M(f, m)$  に変換し、

前記周波数領域の信号値  $X_1(f, m), \dots, X_M(f, m)$  からなる第1のベクトル  $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$  を、周波数  $f$  ごとに  $N$  個ずつのクラスタ  $C_i(f)$  ( $i = 1, \dots, N$ ) にクラスタリングし、

前記各クラスタ  $C_i(f)$  を代表する第2のベクトル  $a_i(f)$  を算出し、前記第2のベクトル  $a_i(f)$  から  $V$  ( $1 \leq V \leq M$ ) 個の第3のベクトル  $a_p(f)$  ( $p = 1, \dots, V$ ) を抽出し、

前記第3のベクトル  $a_p(f)$  の集合を  $G_k$  とし、 $G_k^c$  を  $G_k$  の補集合とし、 $D(\alpha, \beta)$  をベクトル  $\alpha$  と  $\beta$  とのマハラノビス平方距離とした場合における、

[数61]

$$M(f, m) = \begin{cases} 1 & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

で示されるマスク  $M(f, m)$  を生成し、

前記マスク  $M(f, m)$  と前記第1のベクトル  $X(f, m)$  との積を演算し、 $V$  個の前記信号源から発せられた信号の値を抽出する、処理を実行するプロセッサと、

を有することを特徴とする信号分離装置。

- [24]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離装置であって、

前記センサにおいて観測された観測信号値  $x_1(t), \dots, x_M(t)$  を格納する記憶部

と、

前記記憶部に接続され、

前記観測信号値 $x_1(t), \dots, x_M(t)$ を周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ に変換し、

前記周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ からなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$ を、周波数 $f$ ごとに $N$ 個ずつのクラスタ $C_i(f)$  ( $i = 1, \dots, N$ )にクラスタリングし、

前記各クラスタ $C_i(f)$ を代表する第2のベクトル $a_i(f)$ を算出し、

前記第2のベクトル $a_i(f)$ から $V$  ( $1 \leq V \leq M$ ) 個の第3のベクトル $a_p(f)$  ( $p = 1, \dots, V$ )を抽出し、

前記第1のベクトル $X(f, m)$ に対し、前記第3のベクトル $a_p(f)$ の集合を $G_k$ とし、 $G_k^c$ を $G_k$ の補集合とし、 $D(\alpha, \beta)$ をベクトル $\alpha$ と $\beta$ とのマハラノビス平方距離とした場合における、

[数62]

$$\max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f))$$

を満たすか否かを判断し、満たすと判断した前記第1のベクトル $X(f, m)$ を、 $V$ 個の前記信号源から発せられた信号の値として抽出する、処理を実行するプロセッサと、を有することを特徴とする信号分離装置。

[25]  $N$  ( $N \geq 2$ ) 個の信号が混合し、 $M$  個のセンサで観測された状況において信号の分離抽出を行う信号分離装置であって、

前記センサにおいて観測された観測信号値 $x_1(t), \dots, x_M(t)$ を格納する記憶部と、

前記記憶部に接続され、

前記観測信号値 $x_1(t), \dots, x_M(t)$ を、周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ に変換し、

前記周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ からなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]^T$ を、周波数 $f$ ごとに $N$ 個ずつのクラスタ $C_i(f)$  ( $i =$

1, ..., N) にクラスタリングし、

前記各クラス  $C_i(f)$  を代表する第2のベクトル  $a_i(f)$  を算出し、

N個の前記第2のベクトル  $a_i(f)$  の中の0個以上の当該第2のベクトルを0ベクトルに置換したM行N列の行列のムーア・ペンローズ型擬似逆行列であるN行M列の分離行列  $W(f, m)$  を算出し、

$Y(f, m) = W(f, m)X(f, m)$  の演算により、分離信号ベクトル  $Y(f, m) = [Y_1(f, m), \dots, Y_N(f, m)]^T$  を算出する、処理を実行するプロセッサと、  
を有することを特徴とする信号分離装置。

[26] M個のセンサで観測されたN( $N \geq 2$ )個の信号の混合である観測信号の値を周波数領域の信号値に変換する手順と、

前記周波数領域の信号値を用い、前記センサ間における観測値の相対値(相対値の写像も含む)を、各周波数において算出する手順と、

前記相対値をN個のクラスにクラスタリングする手順と、

前記の各クラスの代表値を算出する手順と、

前記代表値を用い、前記周波数領域の信号値から、 $V(2 \leq V \leq M)$  個の信号源から発せられた信号から成る混合信号の値を抽出するためのマスクを作成する手順と、

前記マスクを用い、前記周波数領域の信号値から前記混合信号の値を抽出する手順と、

前記混合信号の値から、V個の信号の値を分離抽出する手順と、

をコンピュータに実行させるための信号分離プログラム。

[27] M個のセンサで観測されたN( $N \geq 2$ )個の信号の混合である観測信号の値を周波数領域の信号値に変換する手順と、

前記周波数領域の信号値を用い、センサ間における観測値の相対値(相対値の写像も含む)を各周波数において算出する手順と、

前記相対値をN個のクラスにクラスタリングする手順と、

前記の各クラスの代表値を算出する手順と、

1個の前記代表値を含む所定の範囲内にある前記相対値に対してハイレベル値をとり、前記所定の範囲内でない前記代表値に対してローレベル値をとり、前記相対値



の変化に伴う当該ハイレベル値から当該ローレベル値への推移が連続的な関数であるマスクを作成する手順と、

前記周波数領域の信号値に前記マスクを乗じ、1個の信号源から発せられた信号の値を抽出する手順と、

をコンピュータに実行させるための信号分離プログラム。

- [28] M個のセンサで観測されたN( $N \geq 2$ )個の信号の混合である観測信号の値 $x_1(t), \dots, x_M(t)$ を周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ に変換する手順と、  
前記周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ からなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$ を、周波数fごとにN個ずつのクラスタ $C_i(f)$  ( $i = 1, \dots, N$ )にクラスタリングする手順と、

前記各クラスタ $C_i(f)$ を代表する第2のベクトル $a_i(f)$ を算出する手順と、

前記第2のベクトル $a_i(f)$ から $V$  ( $1 \leq V \leq M$ ) 個の第3のベクトル $a_p(f)$  ( $p = 1, \dots, V$ )を抽出する手順と、

前記第3のベクトル $a_p(f)$ の集合を $G_k$ とし、 $G_k^c$ を $G_k$ の補集合とし、 $D(\alpha, \beta)$ をベクトル $\alpha$ と $\beta$ とのマハラノビス平方距離とした場合における、

[数63]

$$M(f, m) = \begin{cases} 1 & \max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f)) \\ 0 & \text{otherwise} \end{cases}$$

で示されるマスク $M(f, m)$ を生成する手順と、

前記マスク $M(f, m)$ と前記第1のベクトル $X(f, m)$ との積を演算し、V個の前記信号源から発せられた信号の値を抽出する手順と、

をコンピュータに実行させるための信号分離プログラム。

- [29] M個のセンサで観測されたN( $N \geq 2$ )個の信号の混合である観測信号の値 $x_1(t), \dots, x_M(t)$ を周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ に変換する手順と、  
前記周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$ からなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]$ を、周波数fごとにN個ずつのクラスタ $C_i(f)$  ( $i = 1, \dots, N$ )にクラスタリングする手順と、

前記各クラスタ $C_i(f)$ を代表する第2のベクトル $a_i(f)$ を算出する手順と、

前記第2のベクトル $a_i(f)$ から $V$  ( $1 \leq V \leq M$ ) 個の第3のベクトル $a_p(f)$  ( $p=1, \dots, V$ ) を抽出する手順と、

前記第1のベクトル $X(f, m)$  に対し、前記第3のベクトル $a_p(f)$  の集合を $G_k$  とし、 $G_k^c$  を $G_k$  の補集合とし、 $D(\alpha, \beta)$  をベクトル $\alpha$  と $\beta$  とのマハラノビス平方距離とした場合における、

[数64]

$$\max_{a_p(f) \in G_k} D(X(f, m), a_p(f)) < \min_{a_q(f) \in G_k^c} D(X(f, m), a_q(f))$$

を満たすか否かを判断し、満たすと判断した前記第1のベクトル $X(f, m)$  を、 $V$  個の前記信号源から発せられた信号の値として抽出する手順と、

をコンピュータに実行させるための信号分離プログラム。

- [30]  $M$  個のセンサで観測された $N$  ( $N \geq 2$ ) 個の信号の混合である観測信号の値 $x_1(t), \dots, x_M(t)$  を、周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$  に変換する手順と、

前記周波数領域の信号値 $X_1(f, m), \dots, X_M(f, m)$  からなる第1のベクトル $X(f, m) = [X_1(f, m), \dots, X_M(f, m)]^T$  を、周波数 $f$  ごとに $N$  個ずつのクラスタ $C_i(f)$  ( $i=1, \dots, N$ ) にクラスタリングする手順と、

前記各クラスタ $C_i(f)$  を代表する第2のベクトル $a_i(f)$  を算出する手順と、

$N$  個の前記第2のベクトル $a_i(f)$  の中の $0$  個以上の当該第2のベクトルを $0$  ベクトルに置換した $M$  行 $N$  列の行列のムーア・ペンローズ型擬似逆行列である $N$  行 $M$  列の分離行列 $W(f, m)$  を算出する手順と、

$Y(f, m) = W(f, m)X(f, m)$  の演算により、分離信号ベクトル $Y(f, m) = [Y_1(f, m), \dots, Y_N(f, m)]^T$  を算出する手順と、

をコンピュータに実行させるための信号分離プログラム。

- [31] 請求項26から30の何れかに記載の信号分離プログラムを格納したコンピュータ読み取り可能な記録媒体。

[図1]

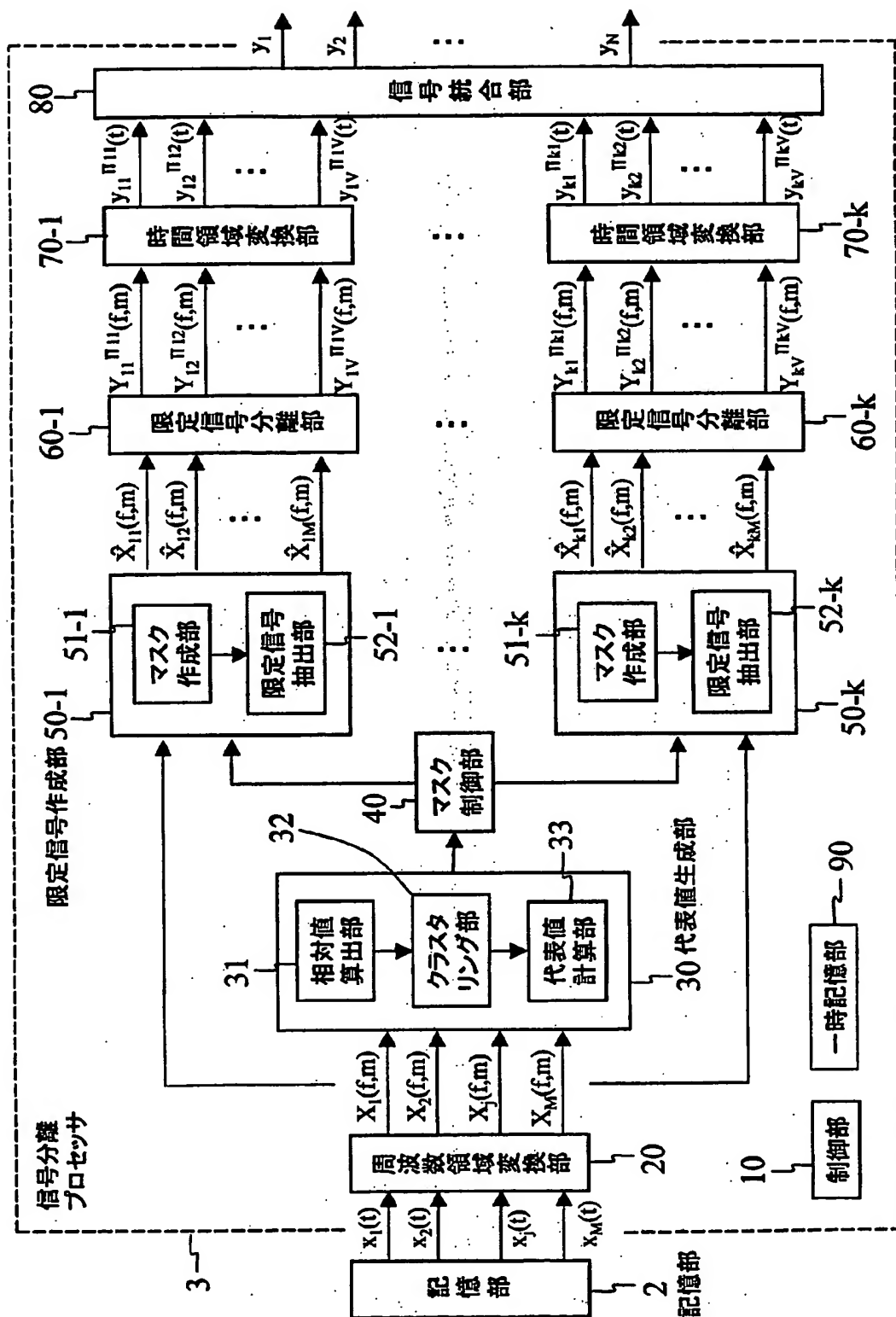


図1

1 信号分離装置

[図2]

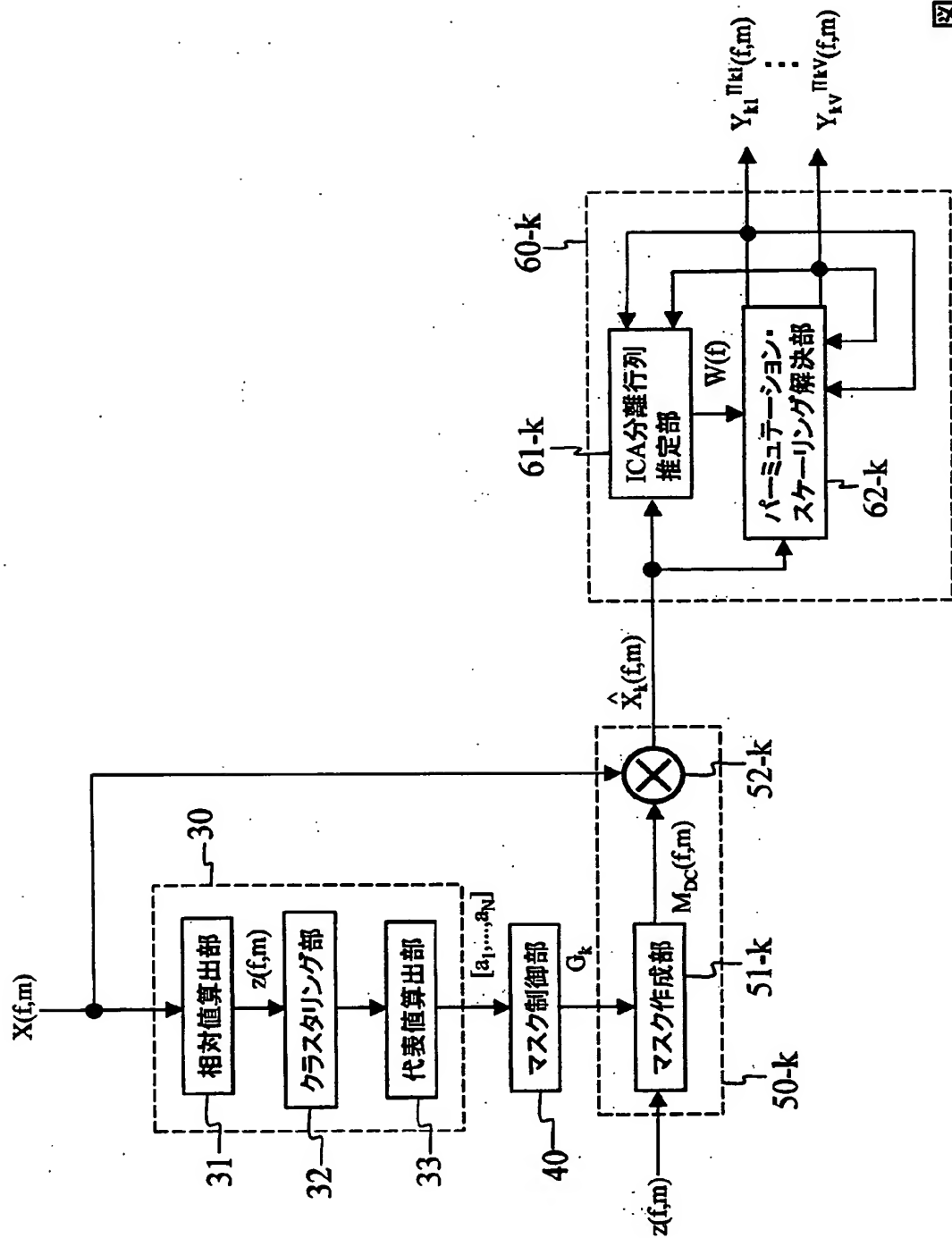


図2

[図3]

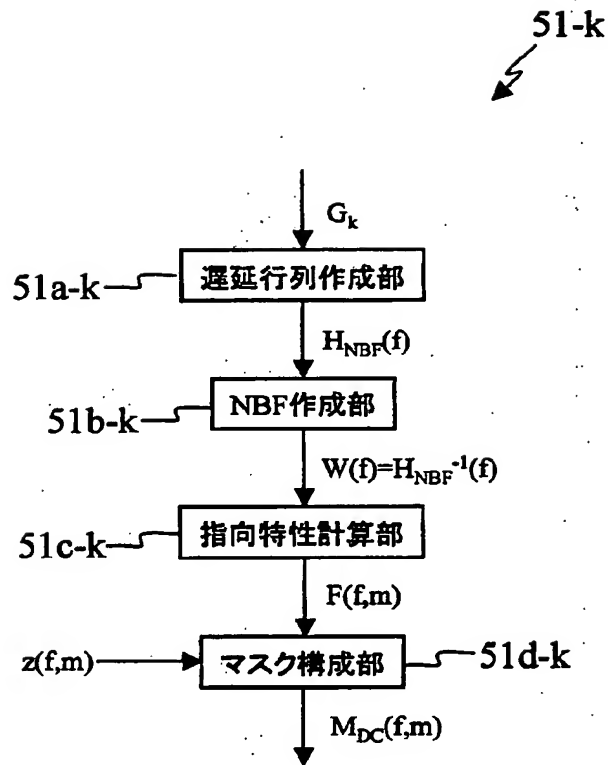


図3

[図4]

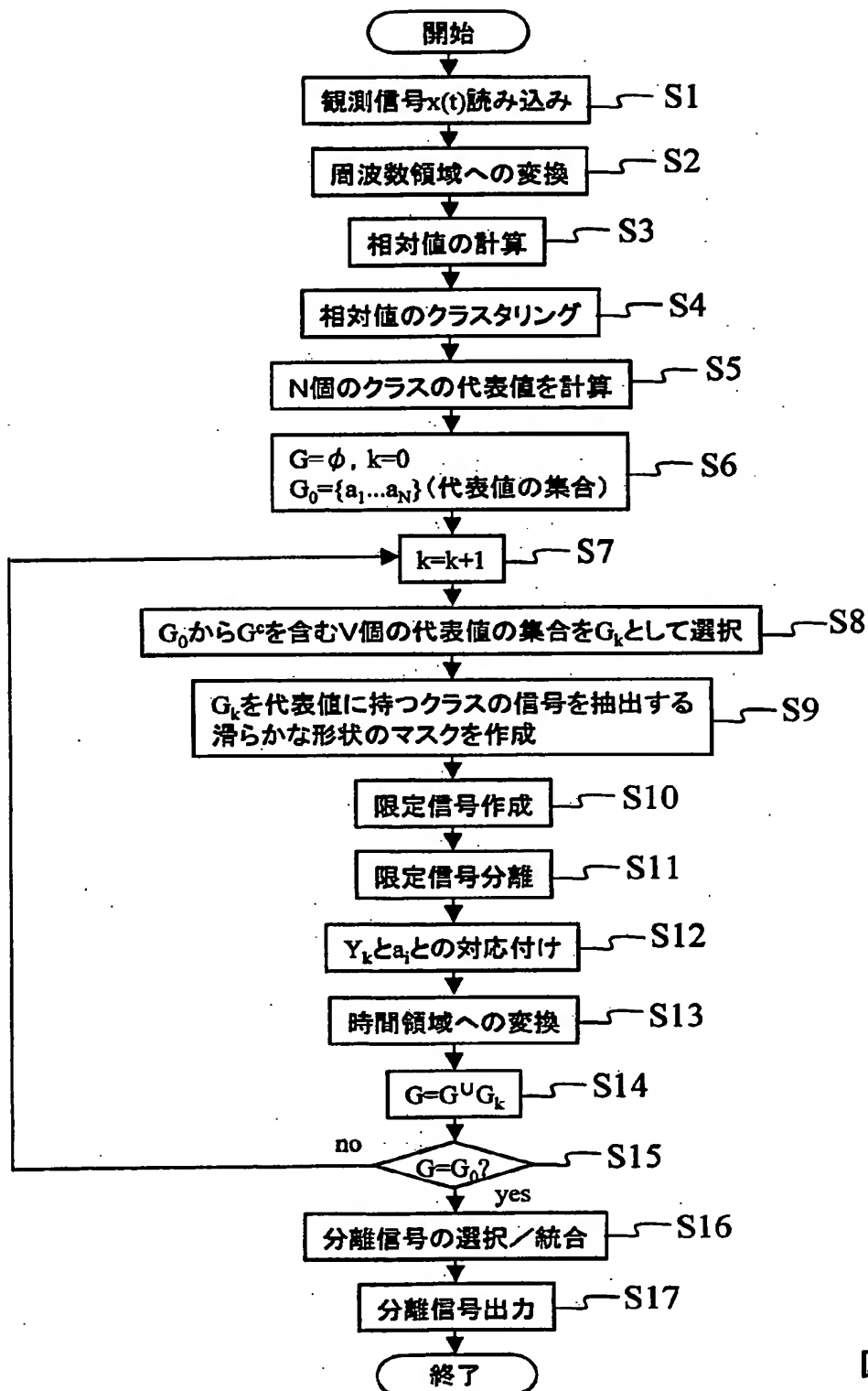


図4

[図5]

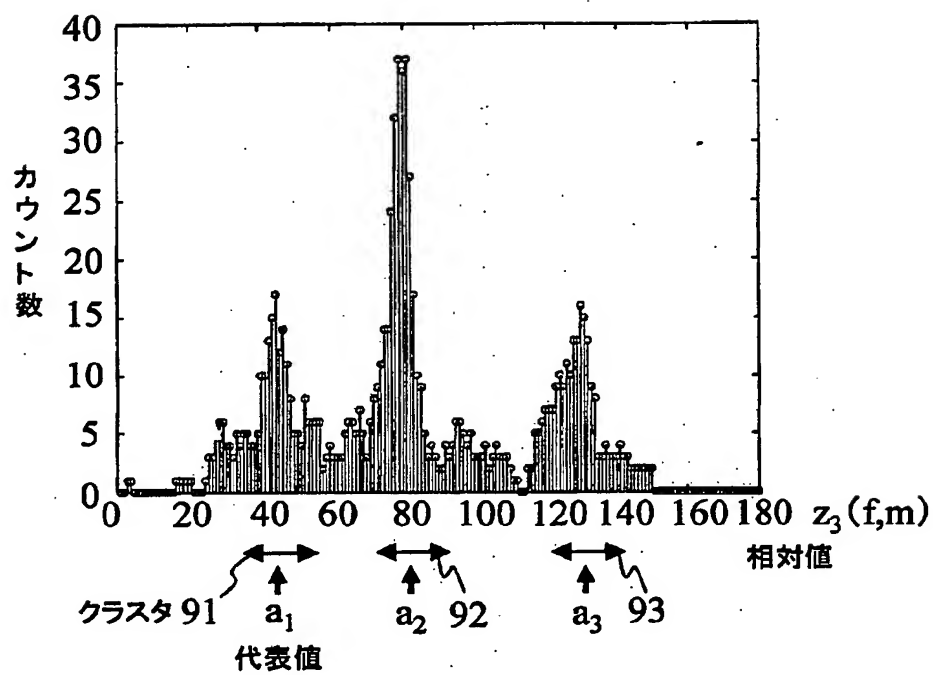


図5

[図6]

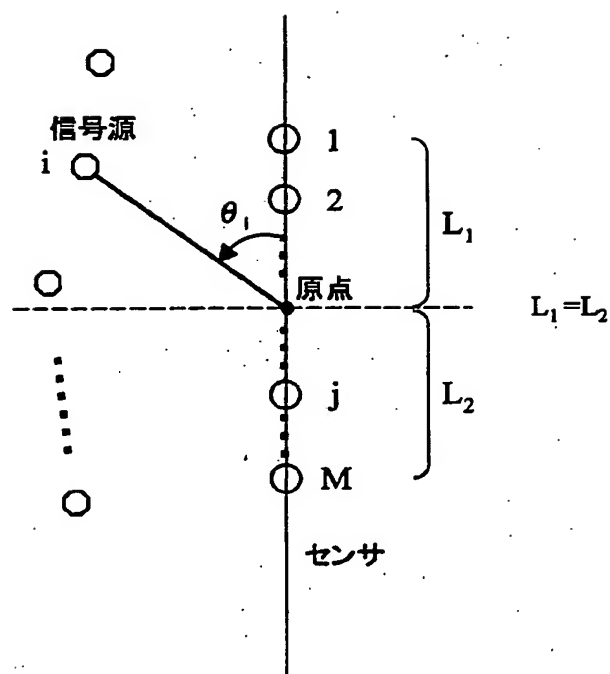


図6



[図7]

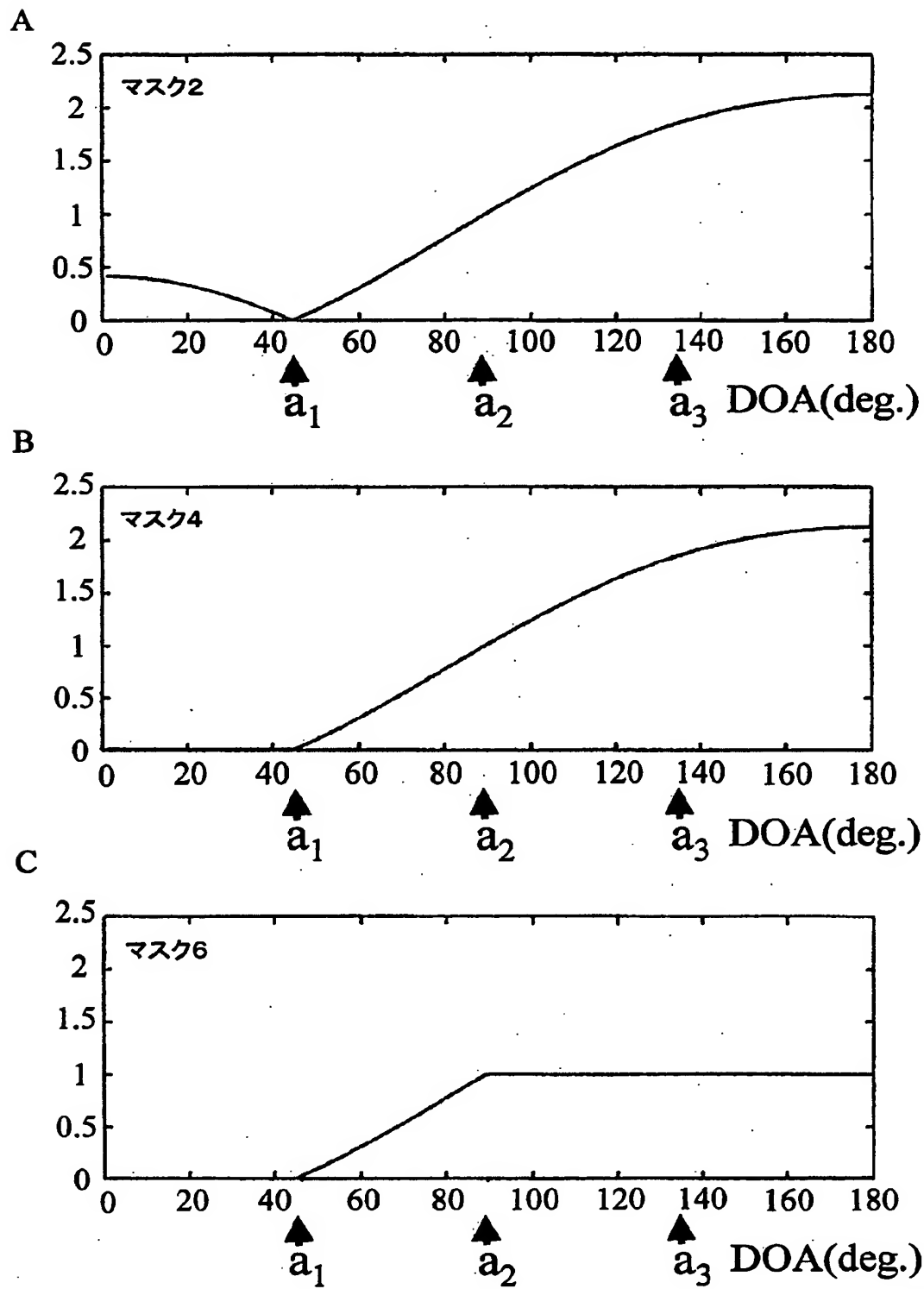
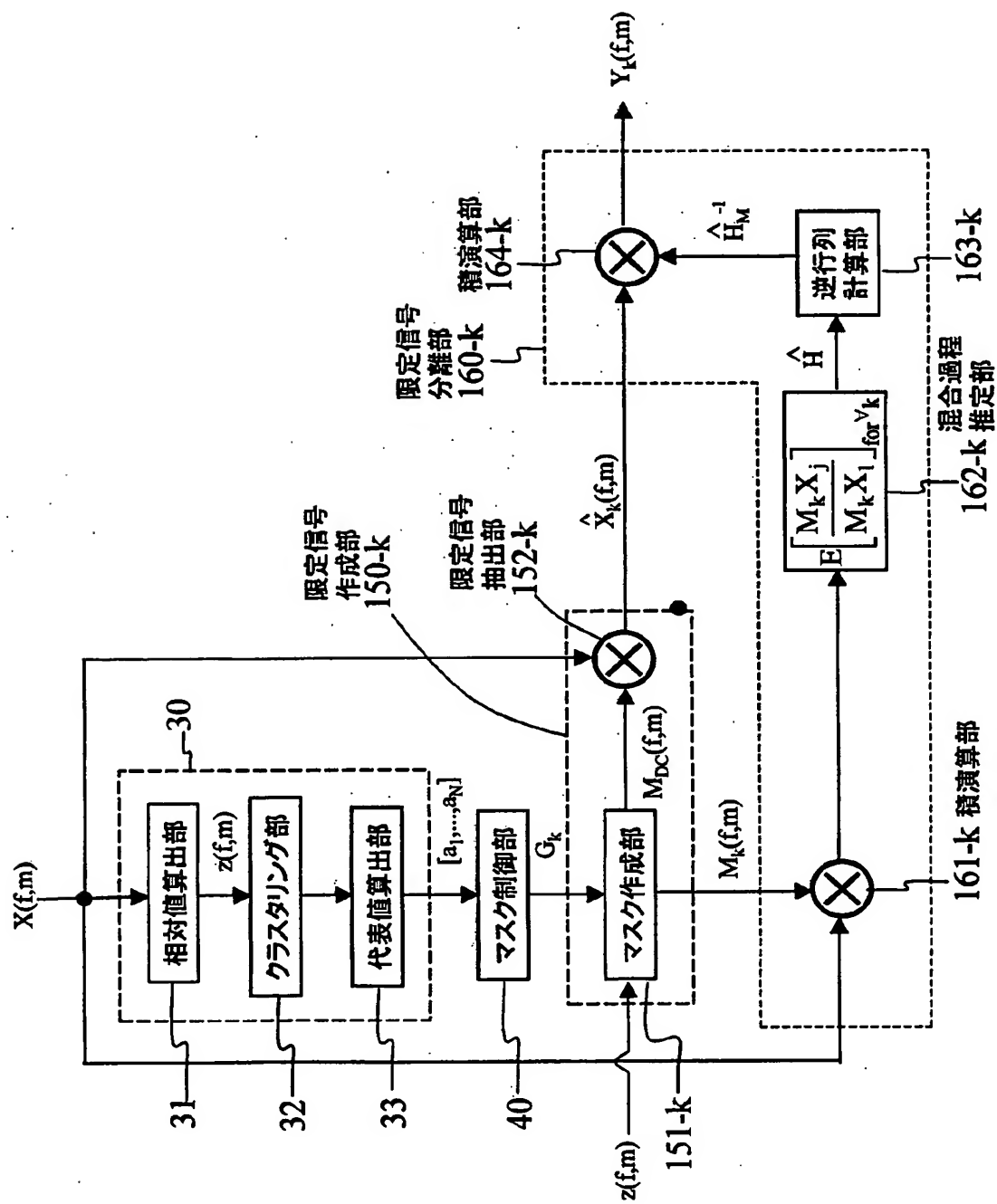


図7

[図8]



[図9]

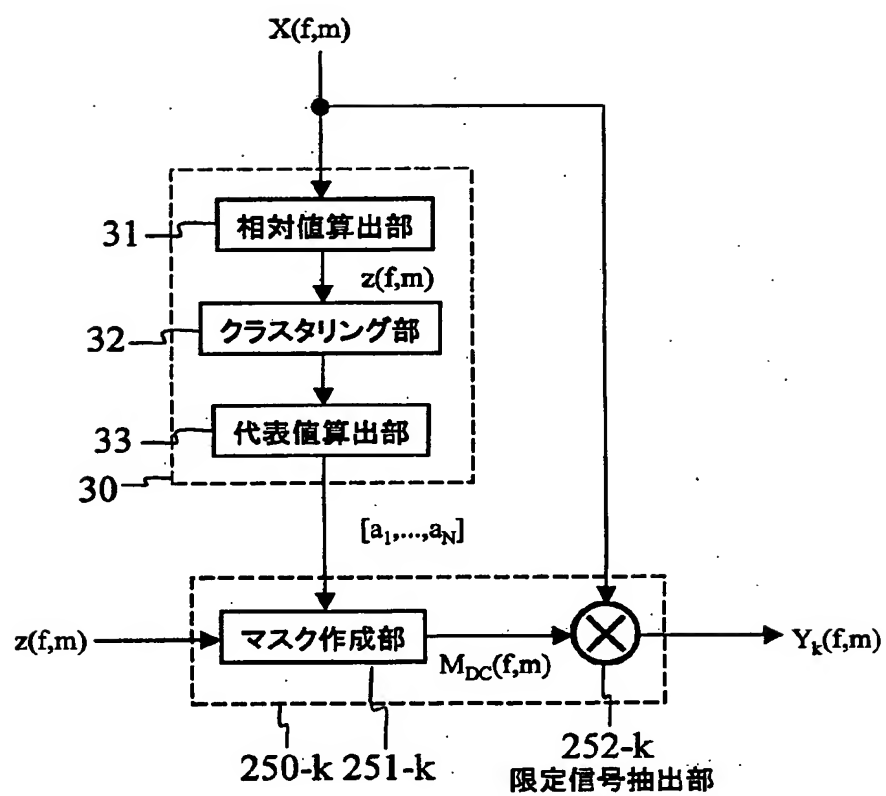


図9

[図10]

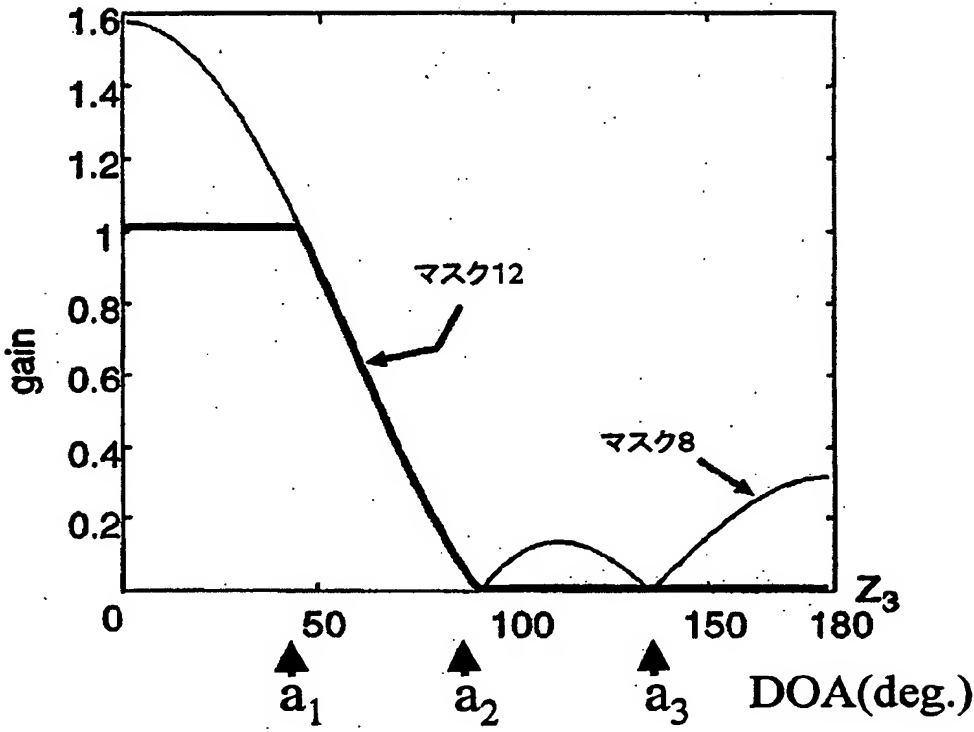


図10

[図11]

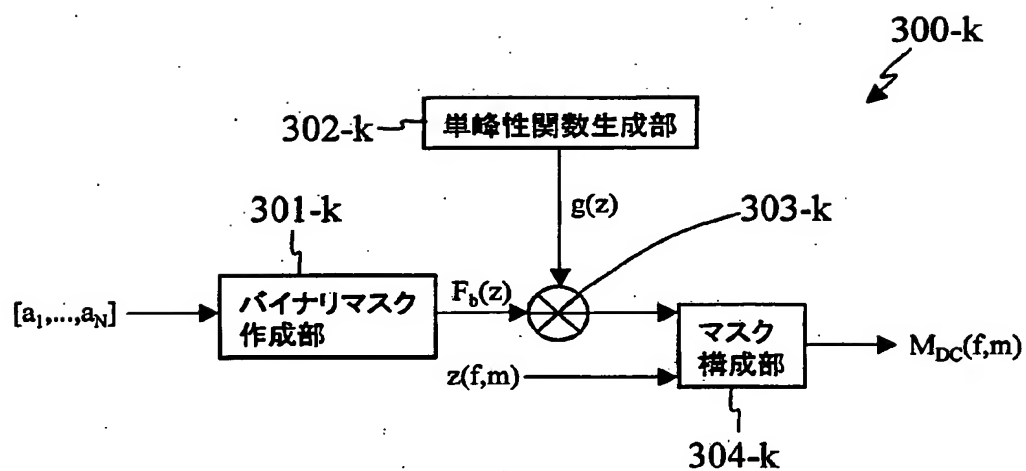


図11

[図12]

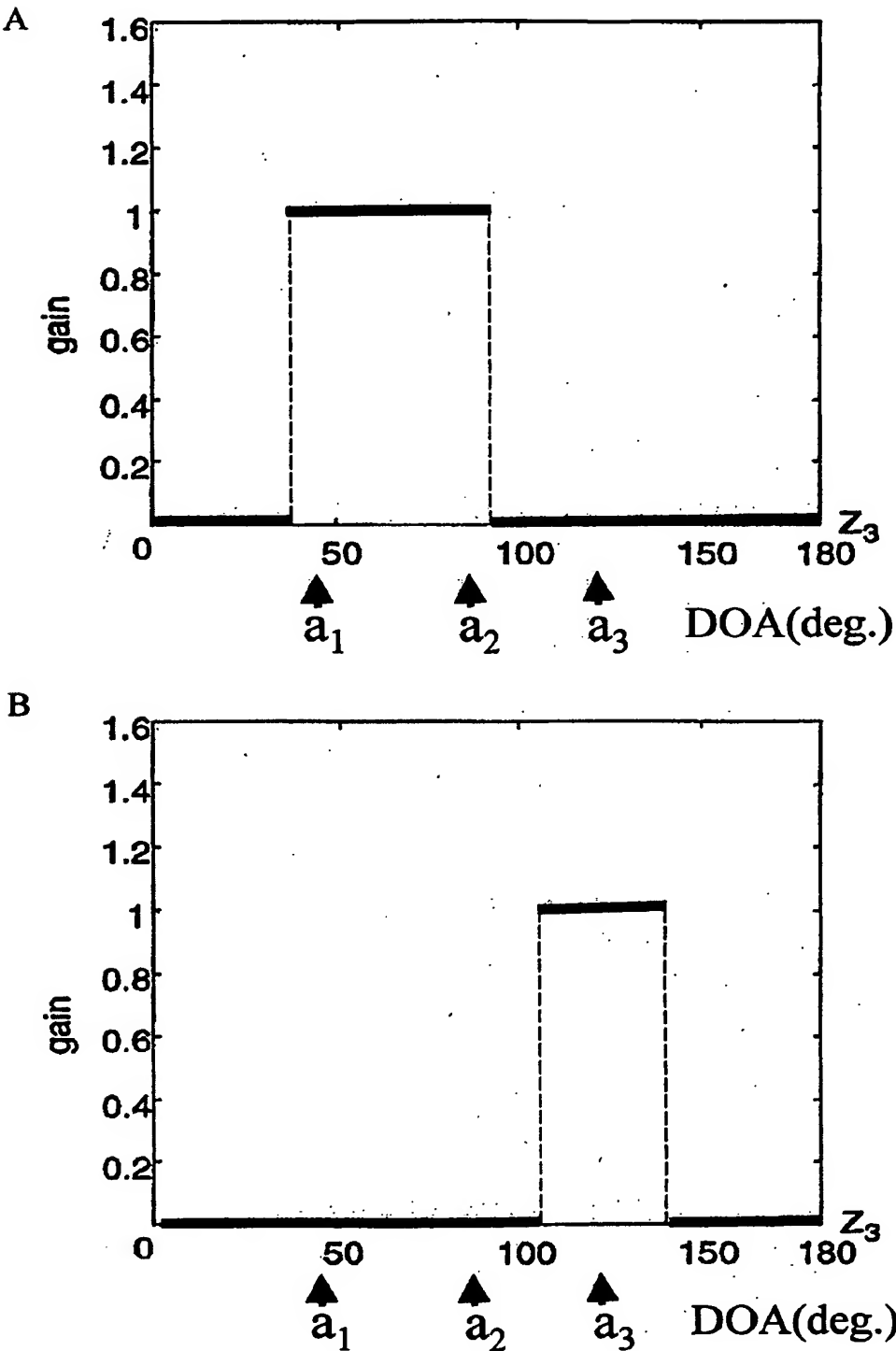


図12

[図13]

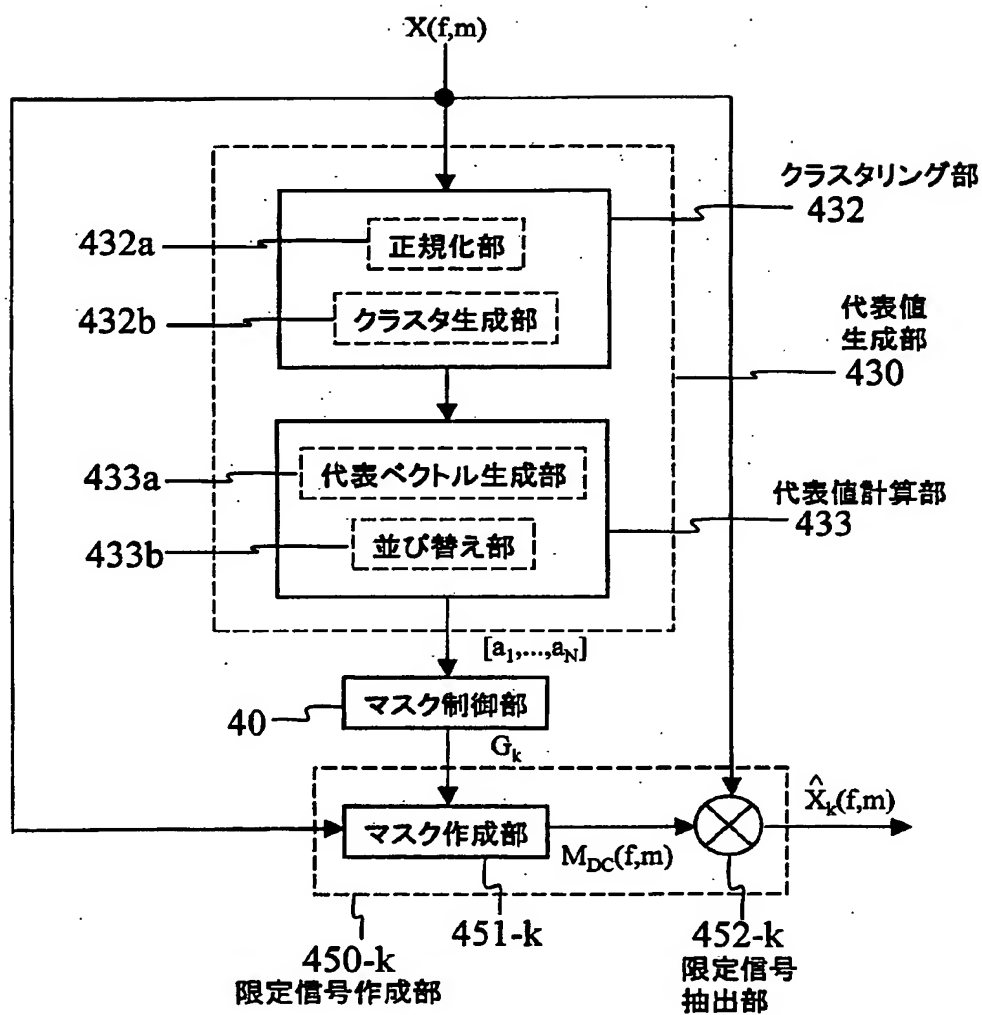


図13

[図14]

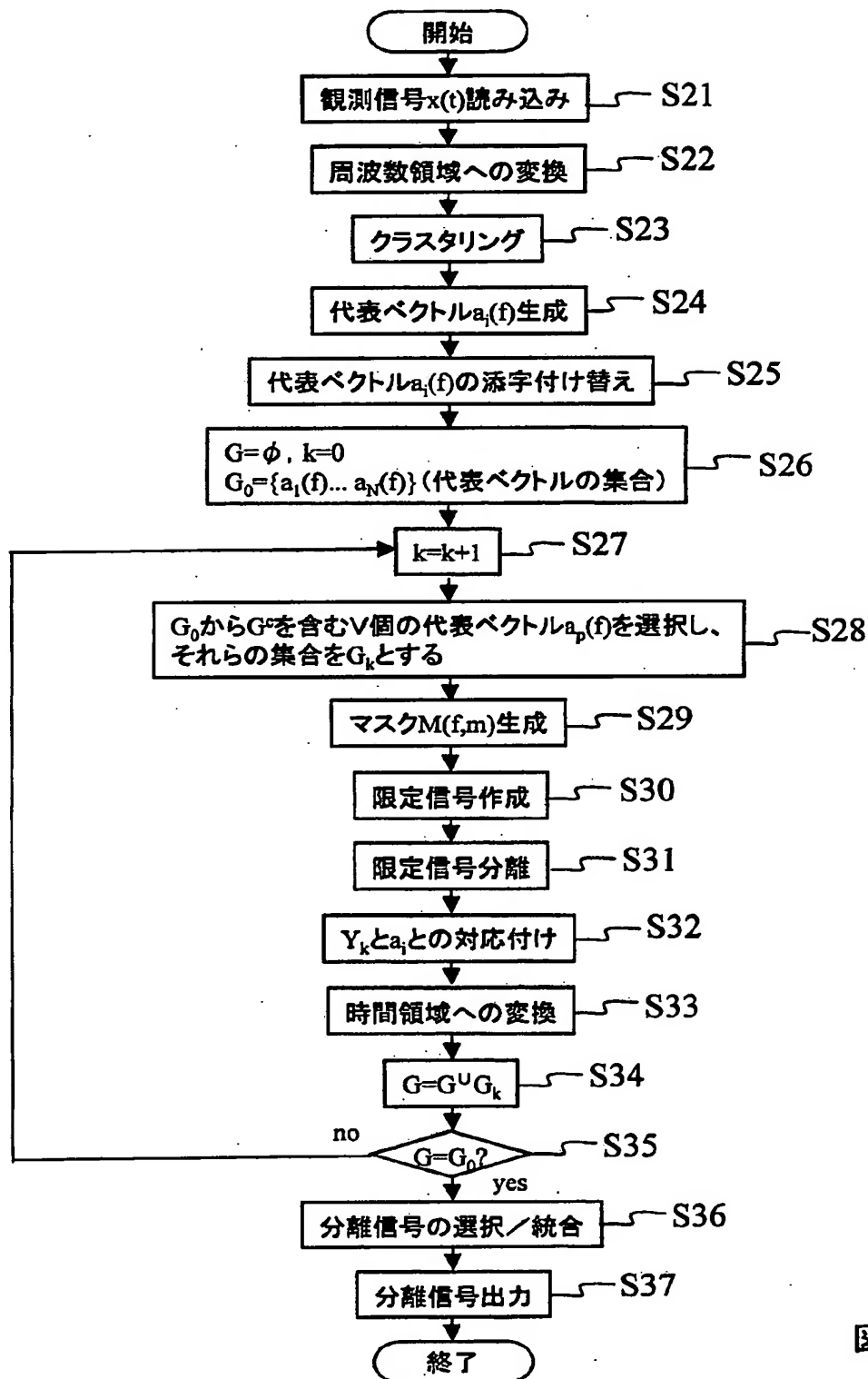


図14



[図15]

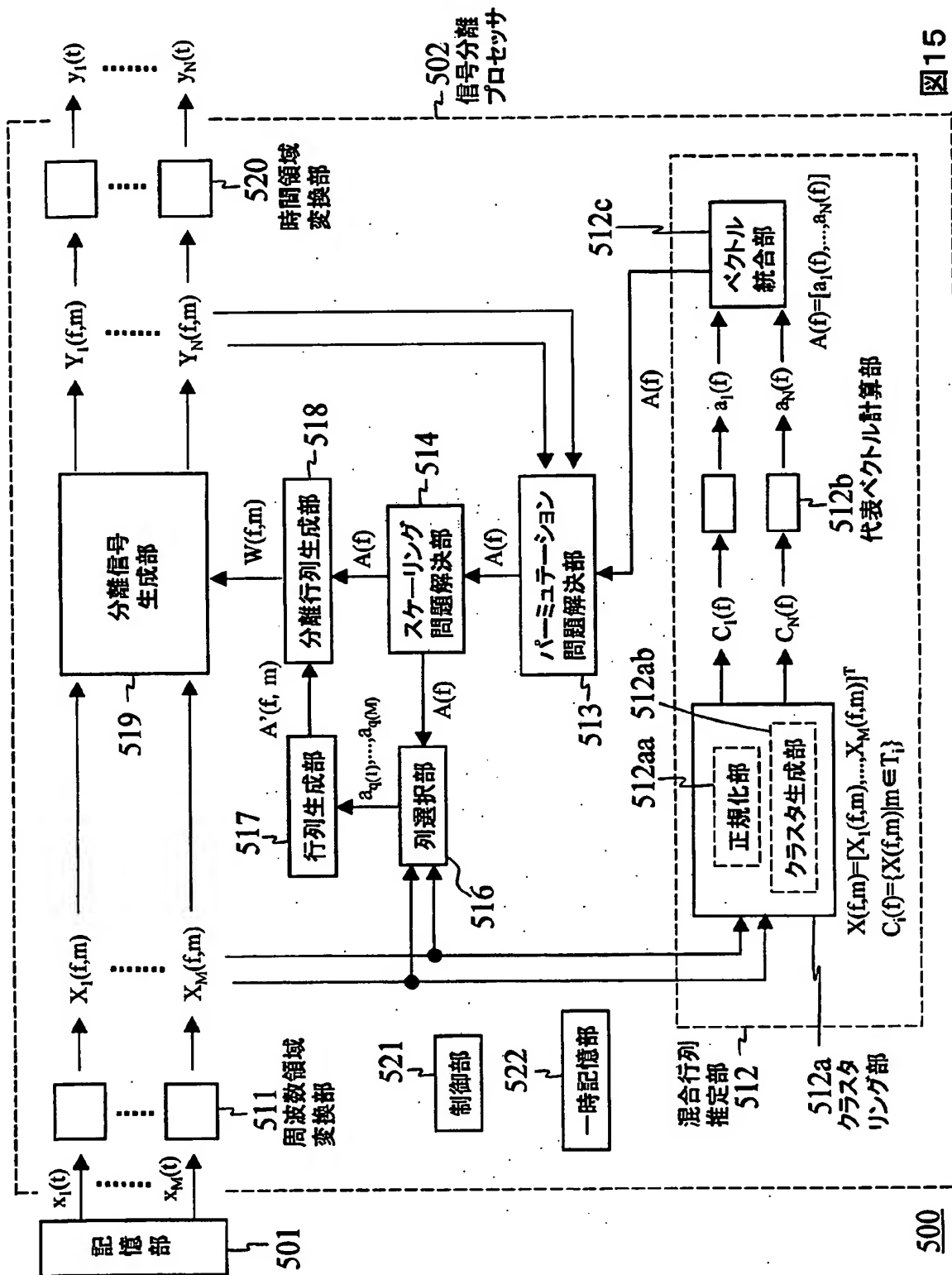


図15

[図16]

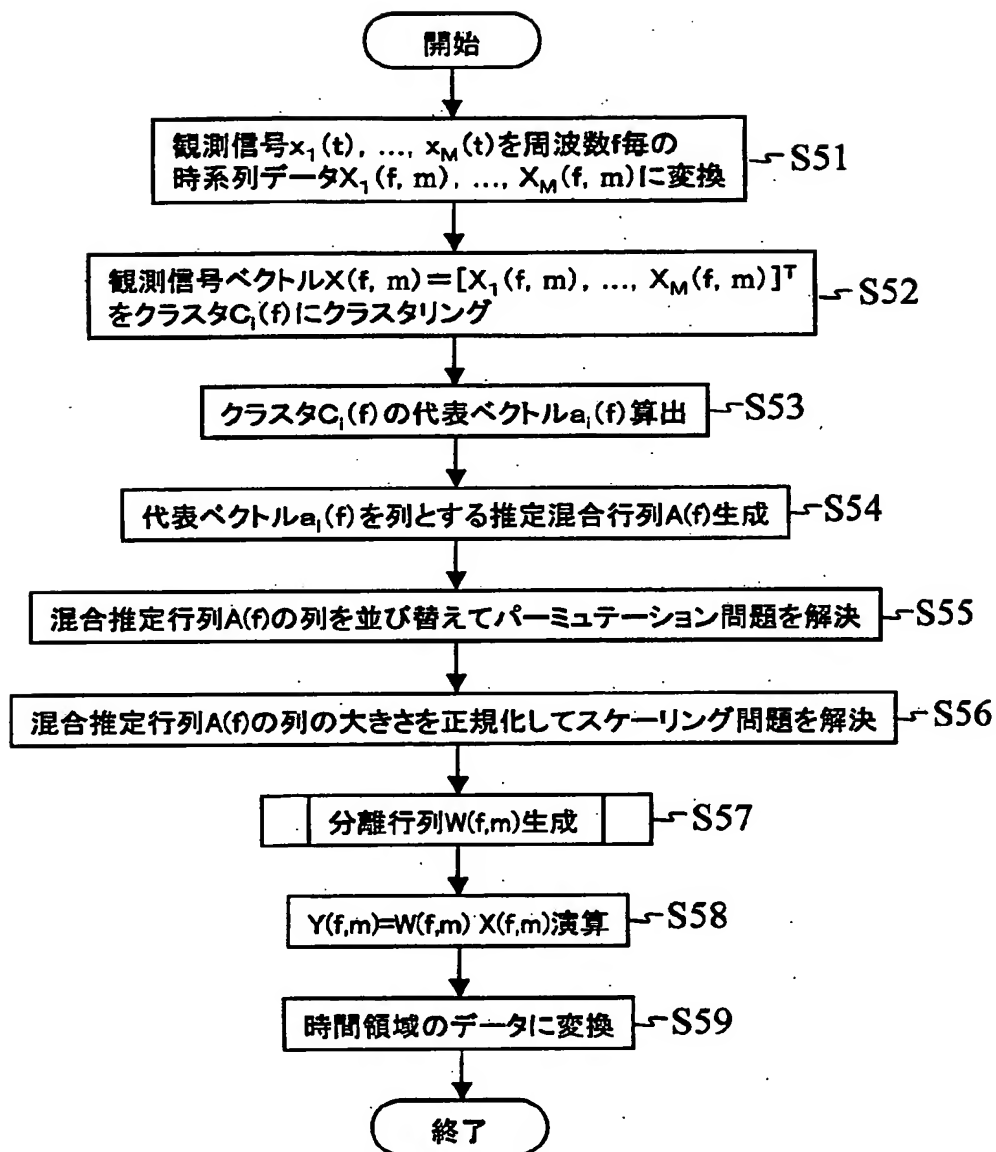


図16

[図17]

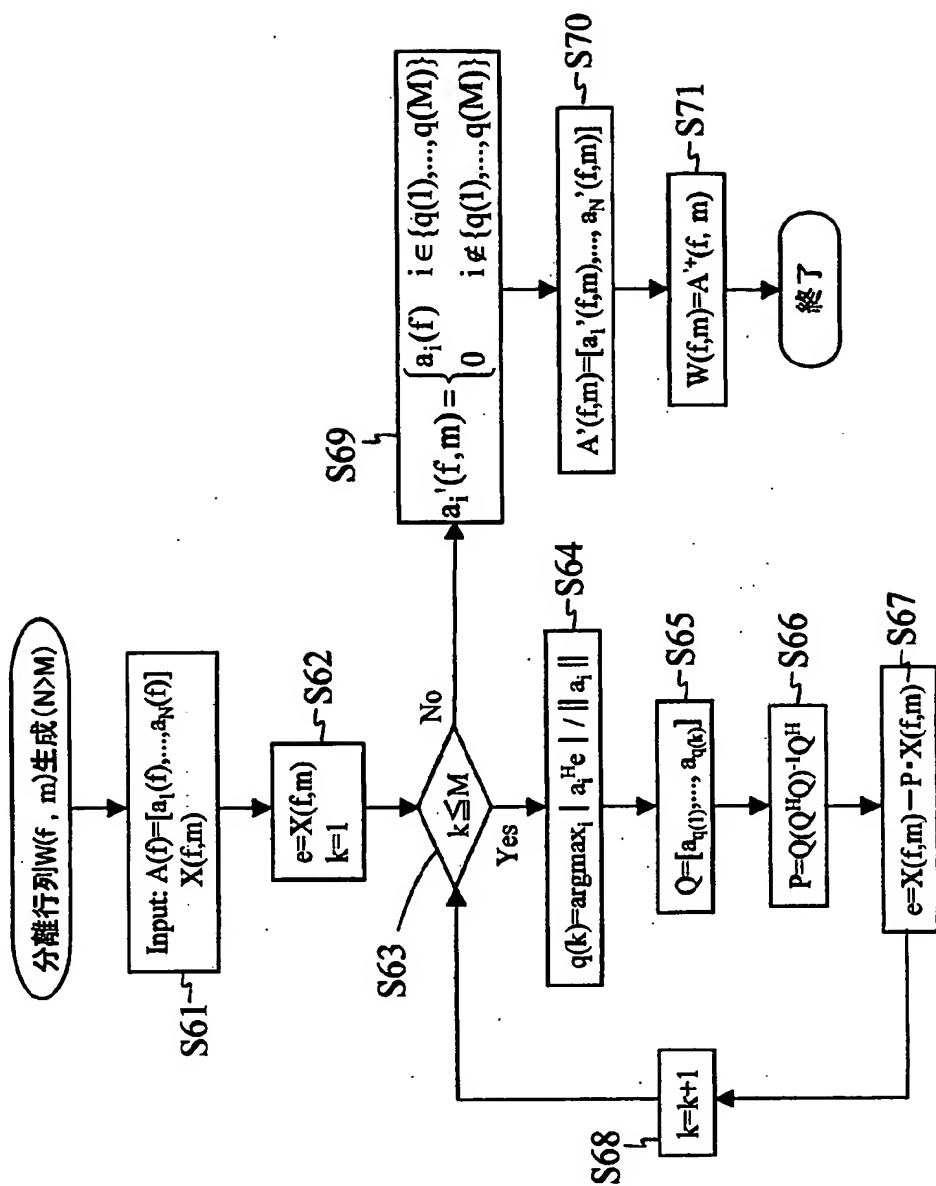


図17

[図18]

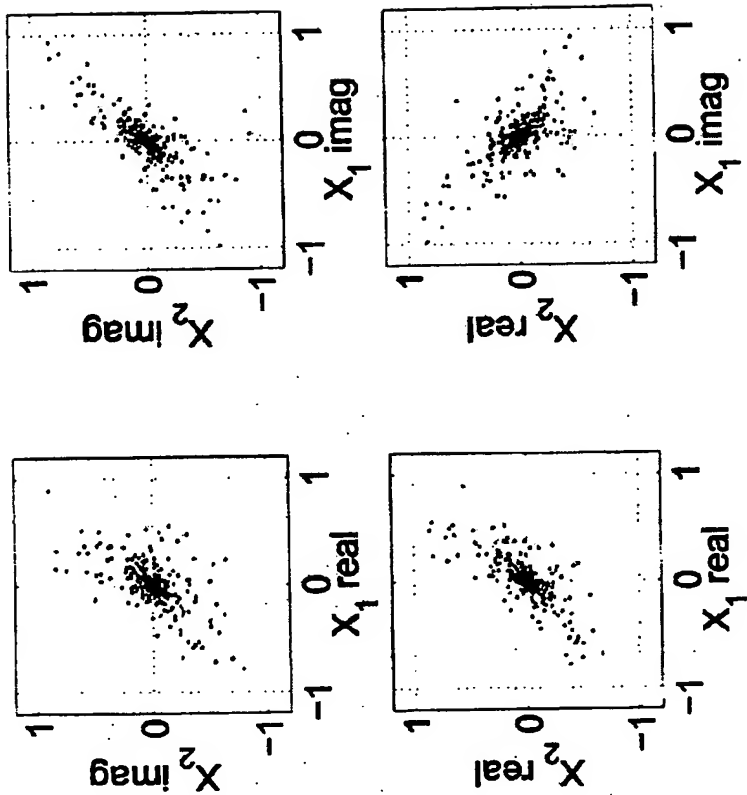


図18

[図19]

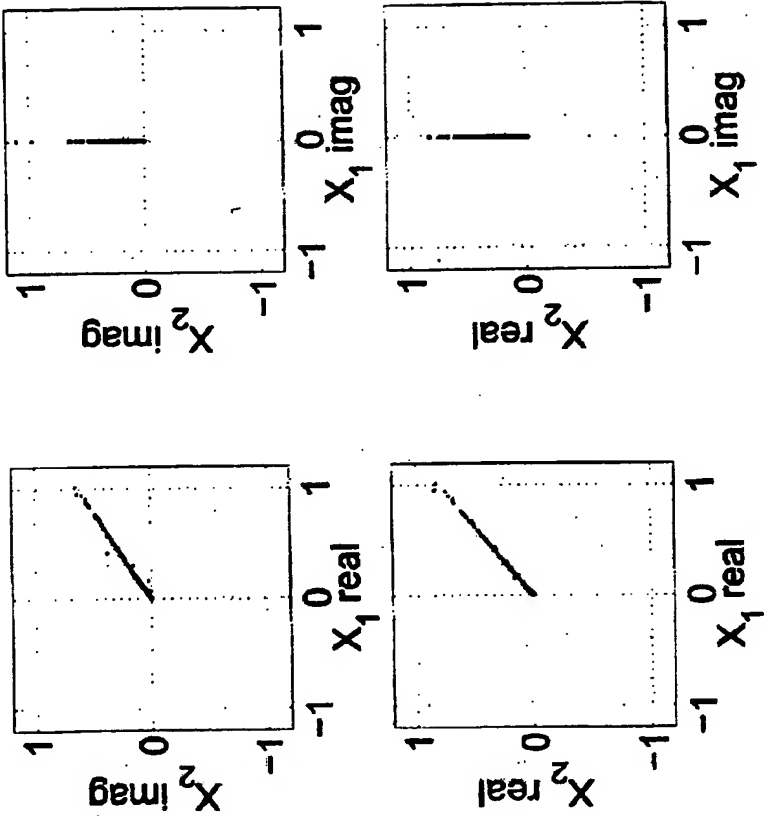


図19

[図20]

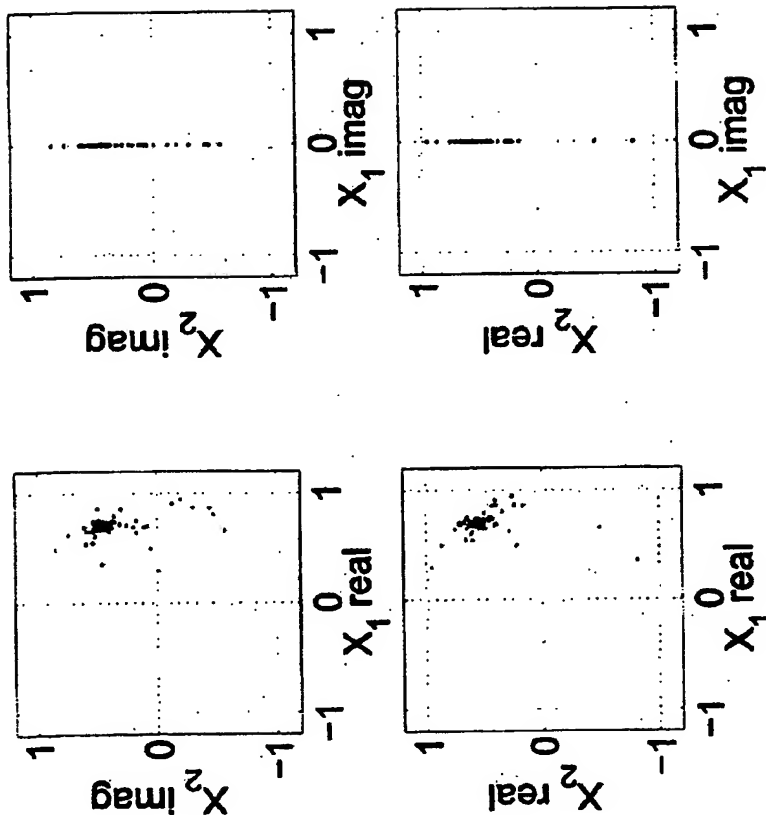


図20

[図21]

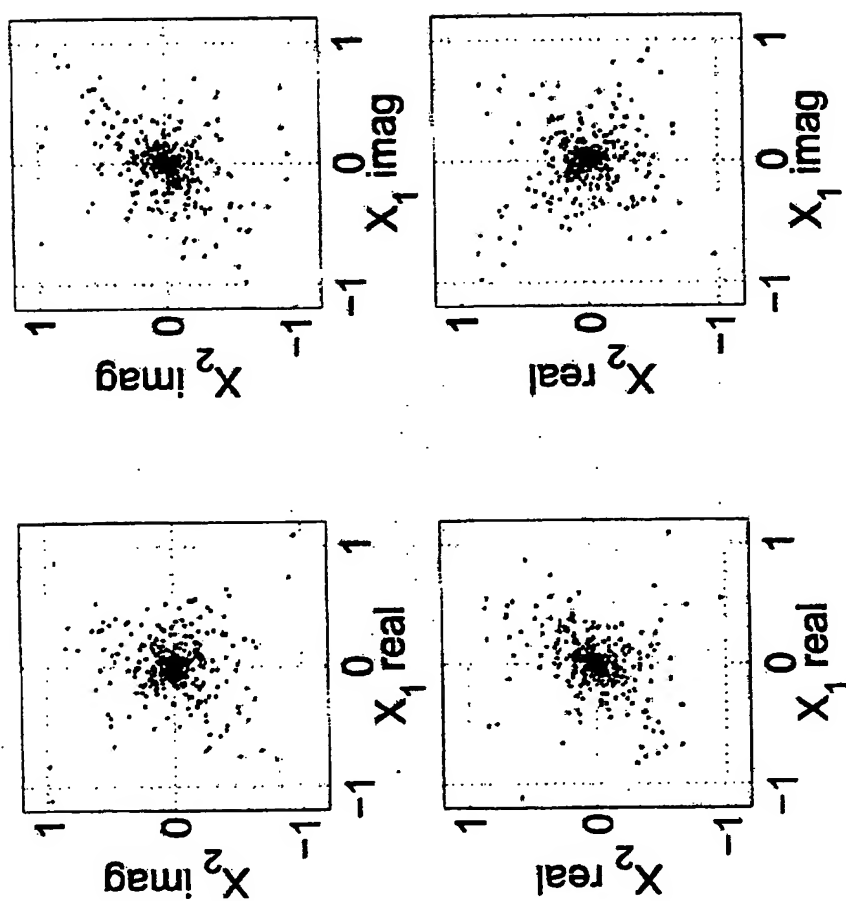


図21

[図22]

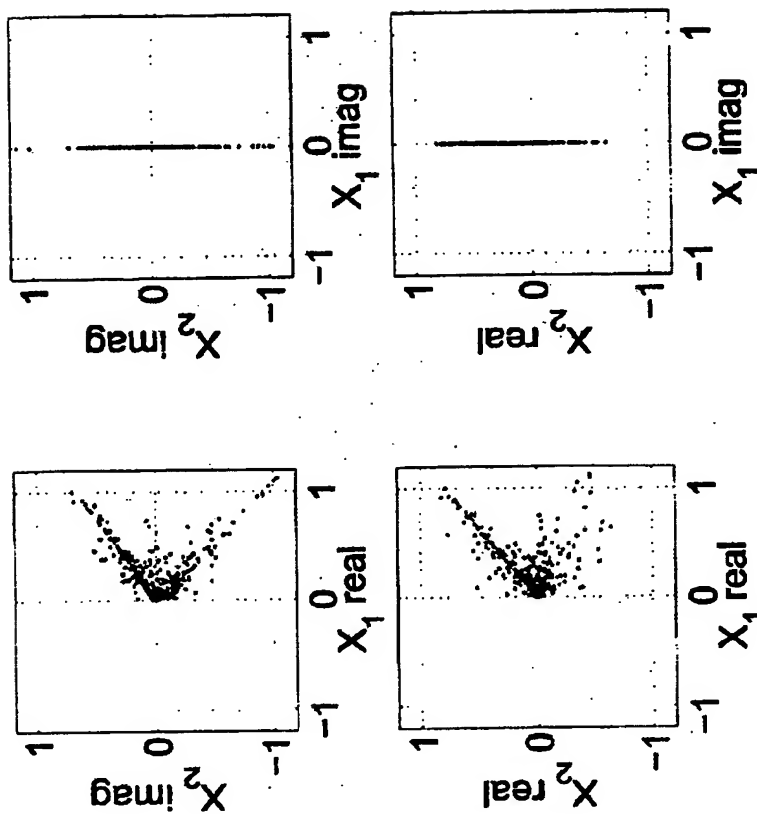


図22



[図23]

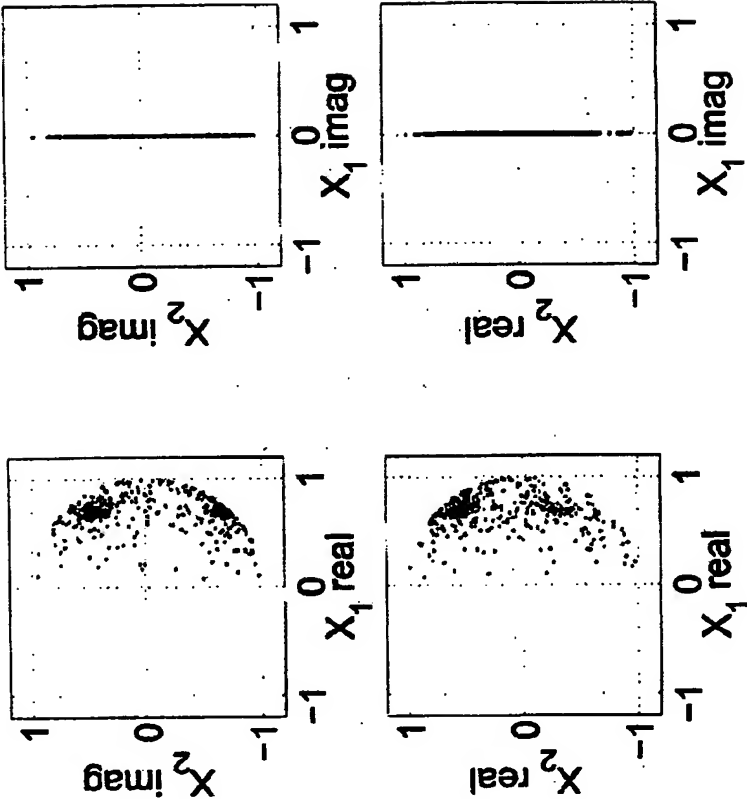


図23

[図24]

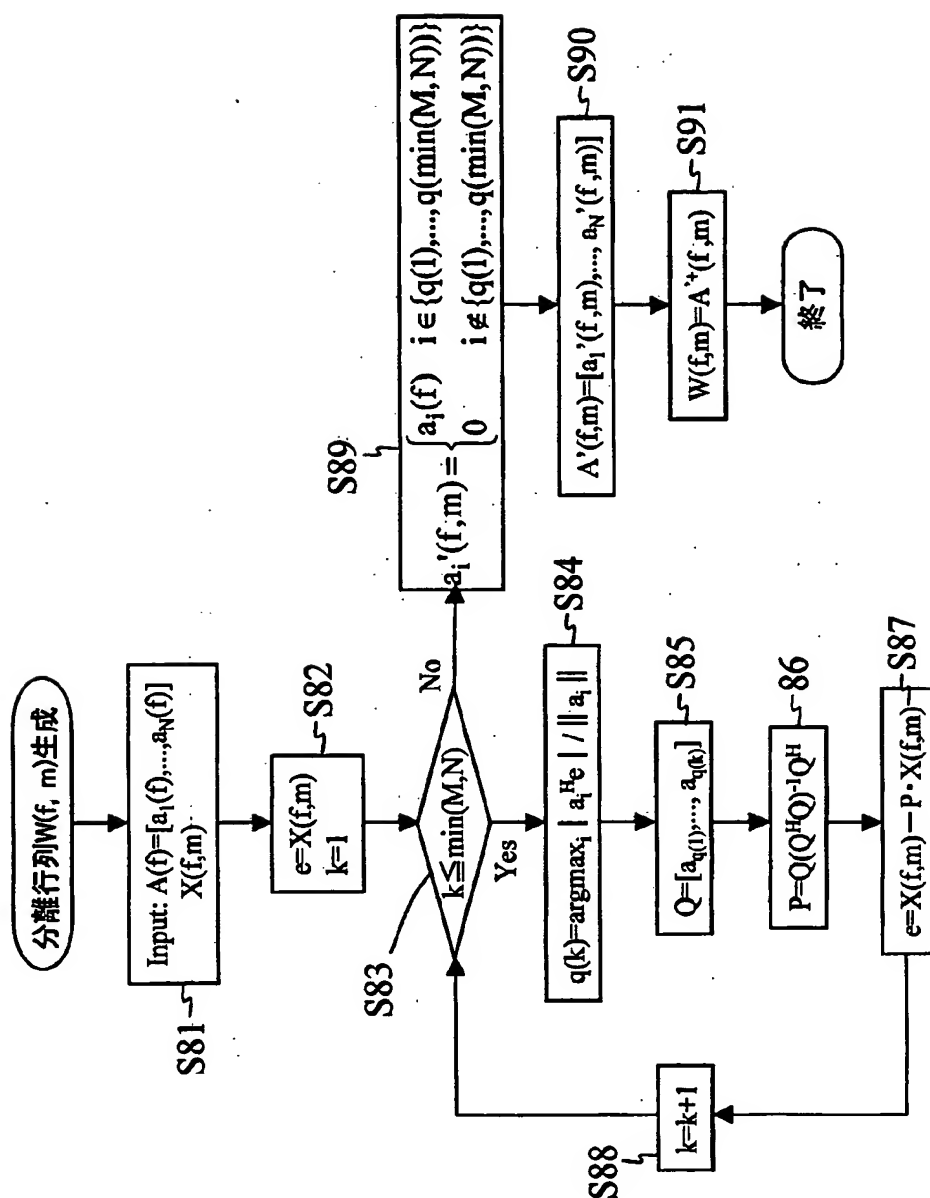


図24

[図25]

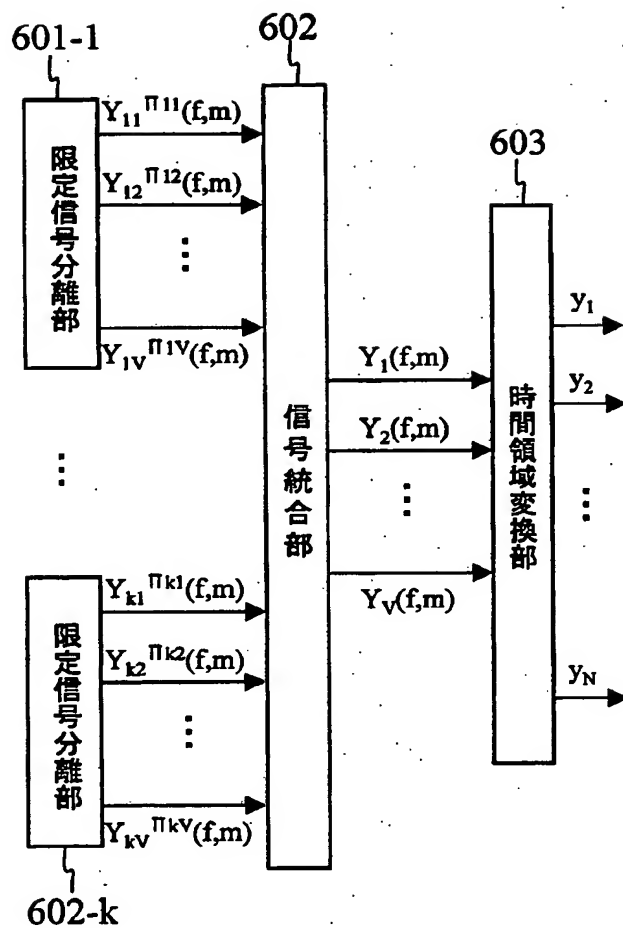


図25

[図26]

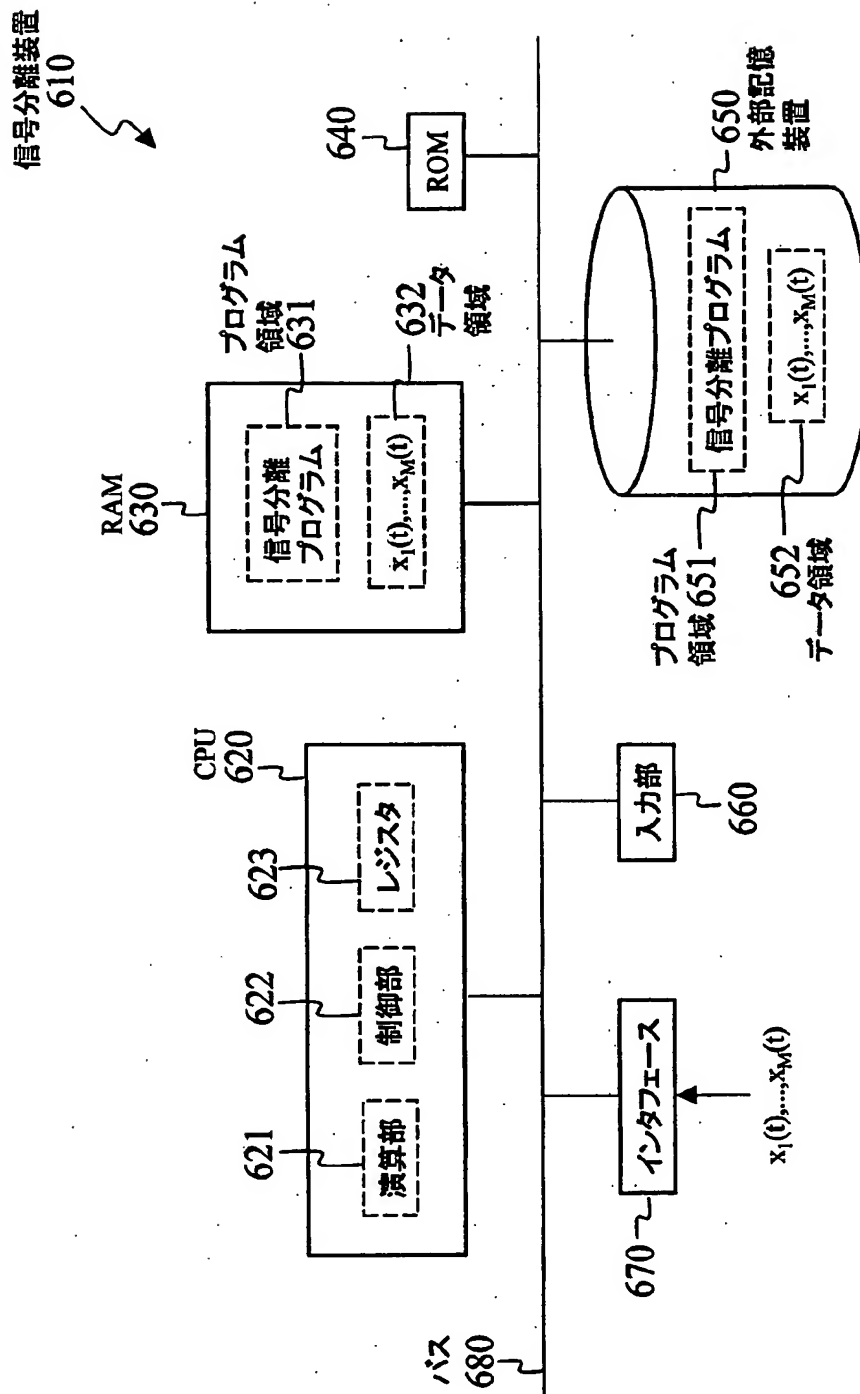
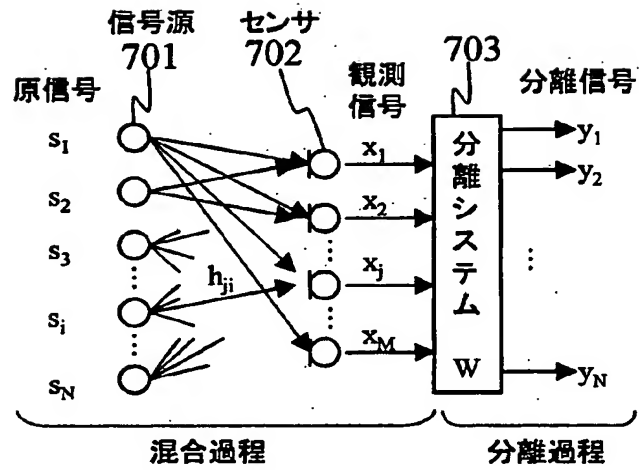


図26

[図27]

A



B

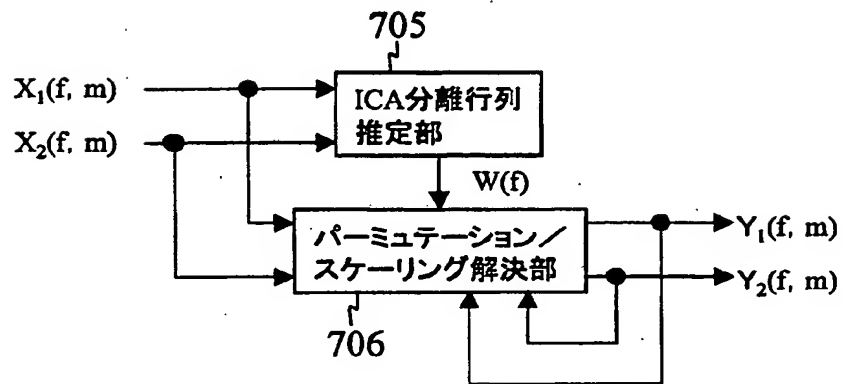
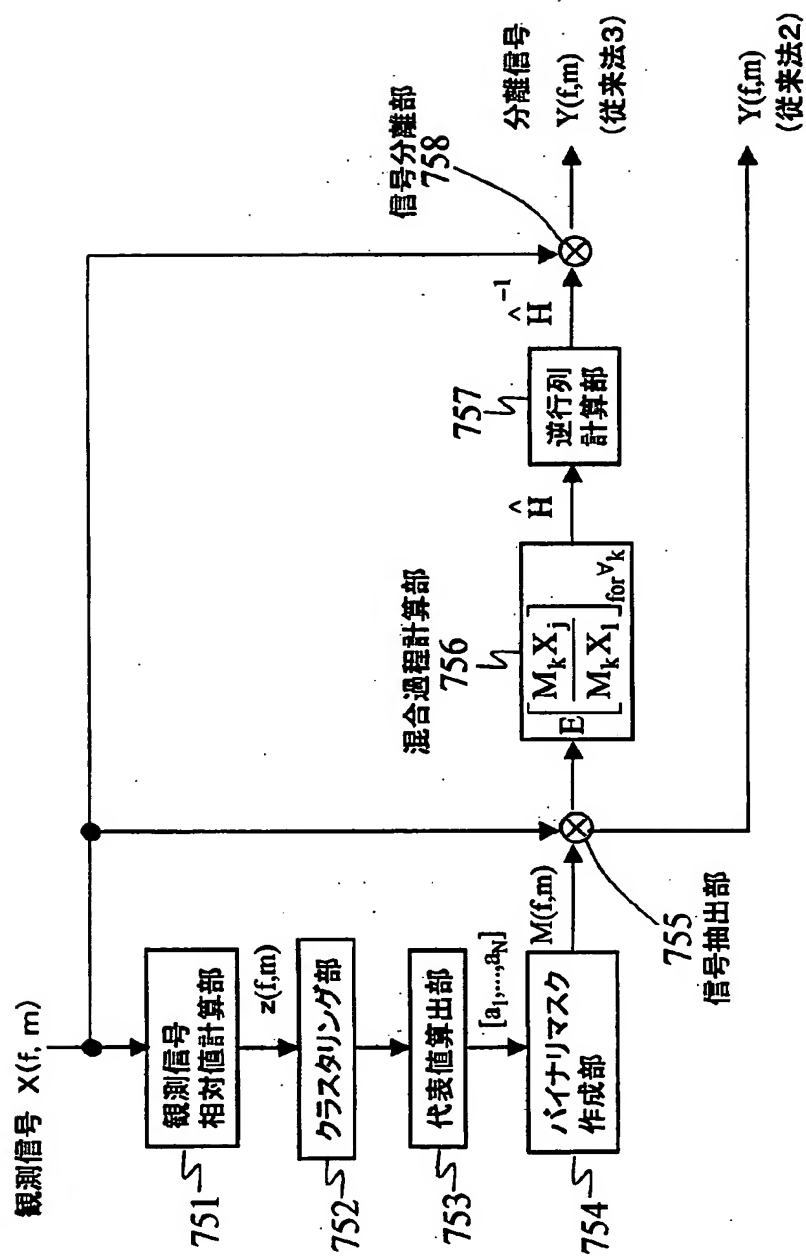


図27

[図28]



28

[図29]

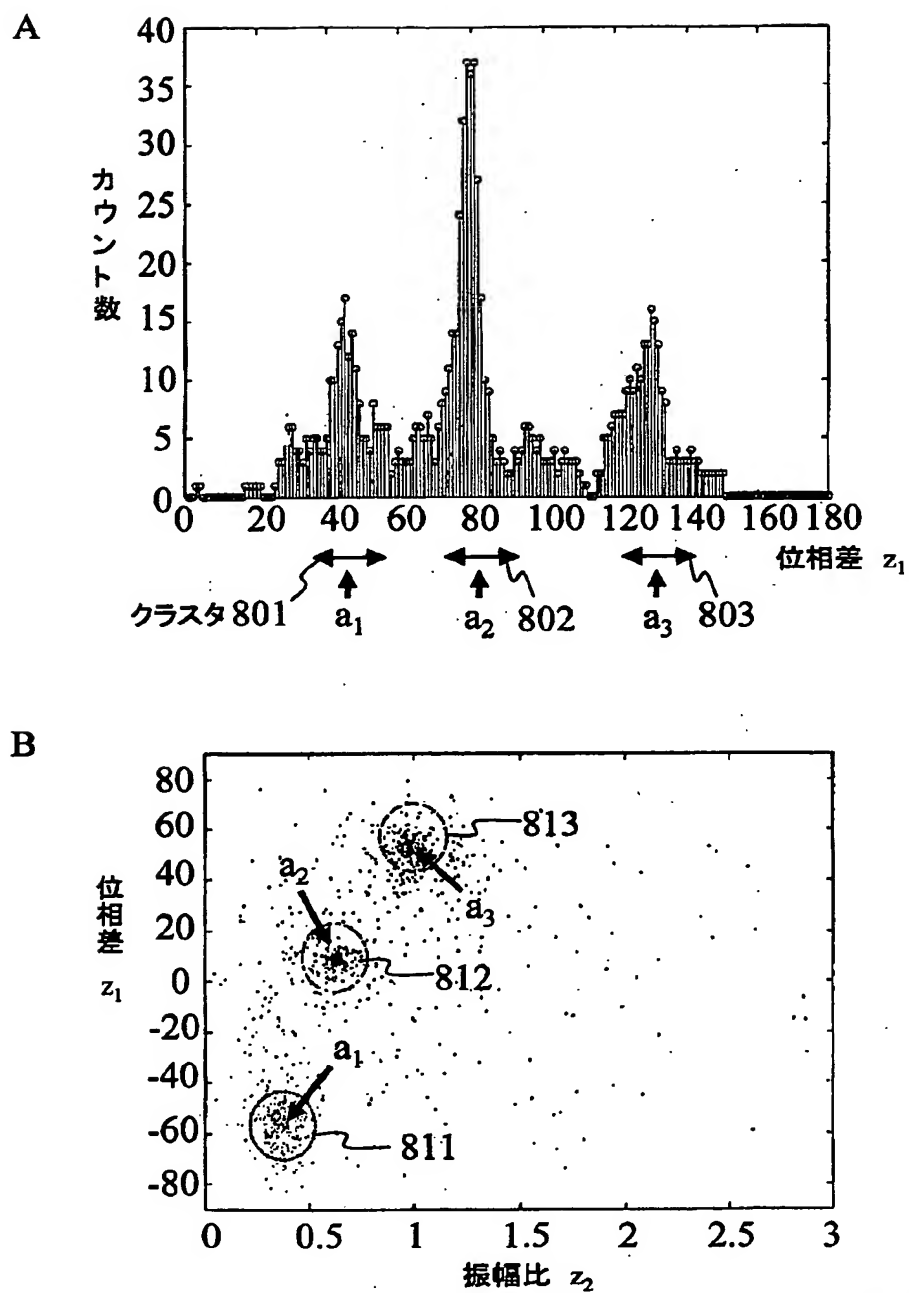


図29

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/012629

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> G10L21/02

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> G10L21/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Jitsuyo Shinan Toroku Koho	1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPlus FILE (JOIS), IEEE Xplore

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	Scott Rickard et al., "On the Approximate W-Disjoint Orthogonality of Speech", Proc. ICASSP, 2002, Vol.1, pages 529 to 532	1-3, 9, 21, 26, 31 4-8, 10-20, 22-25, 27-30
Y A	Hiroshi SARUWATARI, "Onsei· Onkyo Shingo o Taisho toshita Blind Ongen Bunri", The Institute of Electronics, Information and Communication Engineers Gijutsu Kenkyu Hokoku [Tsushin Hoshiki], Vol.101, No.669, 25 February, 2002 (25.02.02), CS2001-134, pages 59 to 66	1-3, 9, 21, 26, 31 4-8, 10-20, 22-25, 27-30

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
26 November, 2004 (26.11.04)Date of mailing of the international search report  
14 December, 2004 (14.12.04)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/012629

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X A	Shoko ARAKI et al., "Jikan Shuhasu Masking to ICA no Heiyo ni yoru Ongensu > Microphone-su no Baai no Blind Ongen Bunri", The Acoustical Society of Japan (ASJ) 2003 Nen Shuki Kenkyu Happyokai Koen Ronbunshu -I-, 17 September, 2003 (17.09.03), 1-P-5, pages 587 to 588	1-3,9,21, 26,31 4-8,10-20, 22-25,27-30

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G10L21/02

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl<sup>7</sup> G10L21/02

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年  
日本国公開実用新案公報 1971-2004年  
日本国登録実用新案公報 1994-2004年  
日本国実用新案登録公報 1996-2004年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JSTPlusファイル (JOIS)  
IEEE Explore

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y A	Scott Rickard et al, 'On the Approximate W-Disjoint Orthogonality of Speech', Proc. ICASSP, 2002, vol.1, p.529-532	1-3, 9, 21, 26, 31 4-8, 10-20, 22-25, 27-30
Y A	猿渡洋, 音声・音響信号を対象としたブラインド音源分離, 電子情報通信学会技術研究報告 [通信方式], Vol. 101, No. 669, 2002. 2. 25, CS2001-134, p. 59-66	1-3, 9, 21, 26, 31 4-8, 10-20, 22-25, 27-30

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの  
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
「O」 口頭による開示、使用、展示等に言及する文献  
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
「&」 同一パテントファミリー文献

国際調査を完了した日

26. 11. 2004

国際調査報告の発送日

14.12.2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

山下 剛史

5C

3148

電話番号 03-3581-1101 内線 3541

## C (続き) . 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
P X A	荒木章子, 他, 時間周波数マスキングと I C A の併用による音源数 > マイク数の 場合のブラインド音源分離, 日本音響学会 2003 年秋季研究発表会講演論文集-I-, 2003. 09. 17, 1-P-5, p. 587-588	1-3, 9, 21, 26, 31 4-8, 10-20, 22-25, 27-30